CONTRIBUTIONS TO ASTRONOMICAL AND MEDICAL INFORMATION PROCESSING

A thesis presented for the degree of
Doctor of Philosophy
in Electrical Engineering
in the
University of Canterbury,
Christchurch, New Zealand,

by

Alastair M. Sinton, B.E.(Hons 1)

May 1986

To my father,

the late

ROY DOUGLAS SINTON

ABSTRACT

A number of specific applications of digital information processing to astronomical speckle imaging and to physiological measurements in medicine are presented in this thesis.

Astronomical speckle processing, a collection of information processing techniques in optical astronomy for high spatial resolution imaging through the Earth's atmosphere, is reviewed in detail.

The performance of simple shift-and-add [Bates and Cady 1980] is studied under conditions of partial isoplanatism and photon limiting. Partial isoplanatism is shown to attenuate and reduce the resolution of detail in the shift-and-add image, to a degree equivalent to the similar effects which occur with speckle interferometry. No spurious fine detail is generated in the shift-and-add image by partial isoplanatism. A variation of shift-and-add is proposed in which any bright, unresolved star serves as a reference object for the formation of a shift-and-add image of fainter detail within the same isoplanatic patch as the star. Computational and astronomical examples are presented demonstrating the potential of this proposal for forming high spatial resolution images of astronomical objects under partially isoplanatic conditions. Photon limiting is shown with the aid of computational examples to cause changes in the form of the shift-and-add image, including an increase of ghosting as the average number of photons per speckle image decreases. The performance of shift-and-add under photon-limited conditions with several different strategies for choosing single "brightest pixels" from speckle images with non-unique brightest pixels is compared.

A new speckle imaging technique called zero-and-add is introduced and described. Zero-and-add compares the complex visibility zeros of individual one-dimensional speckle images to identify the complex visibility zeros, common to all of the speckle images, characterising the object. From these common zeros an image of the object is computed. The effects on zero-and-add of various

types of deviation from isoplanatism in speckle images are investigated.

A computerised system for long-term respiratory monitoring in hospitals is described. The system combines a non-invasive transducer of breathing motions, permitting continuous monitoring of respiratory activity for extended periods in clinical situations, with a general-purpose microcomputer that analyses the breathing motions breath-by-breath. The system has been designed as a tool for the study of the respiratory effects of different techniques of pain relief on hospital patients undergoing surgery. Preliminary clinical results obtained during the development of the system are presented. The system is potentially applicable in a wide variety of clinical and laboratory studies involving respiratory measurements.

A suite of electronic hardware and computerised data processing algorithms for venous occlusion plethysmography [Seagar 1983a, Seagar et al 1984] has been integrated into a self-contained microprocessor-controlled system suitable for routine use in hospitals. The finished system is described. The system automates the acquisition and interpretation of venous occlusion plethysmographic measurements. It allows convenient, non-invasive measurements of limb blood circulation to be made in a wide variety of clinical and medical laboratory settings. Examples of clinical applications of the system in Christchurch hospitals are presented.

ACKNOWLEDGEMENTS

Many people have contributed in greater or lesser part to my Ph.D. research and to the onerous task of preparing this thesis. The list below names individuals and groups who have made important specific contributions to this work. To all of these people, and to the others unnamed here, I acknowledge my debt to you and offer you my thanks for helping to make this work real.

SUPERVISION - Principal supervisor Professor Richard Bates, associate supervisor Professor John Gibbs; who each in their individual ways are largely responsible for the success of this work. Also Dr Peter Gough, who supervised me in optical experiments during 1982 and 1983.

COLLEAGUES AND ADVISORS - Electrical Engineering: Drs Andrew Seagar, Richard Fright, Donald Bailey; Messrs Nigel Brieseman, Bruce Davey, Raveentheran Suntheralingam, Peter Gardenier, Gordon Cameron. Christchurch Clinical School: Miss Adele Wilkinson, Dr Michael Davis.

TECHNICAL SUPPORT - Laurie Hill, Art Vernon.

SECRETARIAL SUPPORT - Margaret Turner, Jan Allchurch, Jenny Palmer.

ASTRONOMICAL DATA - E.K. Hege, J.M. Beckers, J.C. Christou of the Steward Observatory, University of Arizona, Tucson, Arizona, USA. Thanks are also due to N. Baba of the Applied Physics Department, Hokkaido University, Sapporo, Japan and to F. Roddier, National Optical Astronomy Observatories, Tucson, Arizona, USA for supplying astronomical data used by other members of the University of Canterbury speckle group.

FINANCIAL SUPPORT - New Zealand University Grants Committee: Postgraduate Scholarship, Shirtcliffe Fellowship. Christchurch City Council: Christchurch City Council Electricity Department Research Scholarship. National Heart

Foundation of New Zealand: Project Grant No. 322.

FAMILY - I greatly value the care, support, and love of my mother, my father, and my brother Julian throughout my Ph.D. research, and indeed throughout the whole of my life.

TABLE OF CONTENTS

TABLE OF FIGURES AND TABLES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

Figures

CHAPTER 5

Figures

CHAPTER 6

Figures

CHAPTER 7

Figures

CHAPTER 8

Figures

Tables

CHAPTER 9

Figures

Tables

PREFACE

This thesis presents several specific applications of information processing with digital computers, to astronomical speckle imaging and to physiological measurements in medicine. "Information processing" is used here in the sense of image or signal processing, i.e. dealing with measured functions of space (images) or time (signals), rather than the data processing of the commercial world. The disparity between the two areas of application presented here reflects the broad importance of this sort of information processing in applied science, and the excellent opportunities for interdisciplinary research which exist for postgraduate students of electrical engineering at the University of Canterbury.

The thesis is divided into three parts. Part 1 describes my work with the speckle imaging group at the Department of Electrical and Electronic Engineering of the University of Canterbury under the supervision of Professor R.H.T. Bates. Part 2 describes work performed in conjunction with the Division of Anaesthesia of the Christchurch Clinical School of Medicine, under the associate supervision of Professor J.M. Gibbs. Part 3 presents a summary of conclusions and also suggestions for future research. Chapters 1, 2, 3, and 7 are reviews of material relevant to the original research reported in this thesis. Chapters 4, 5, 6, 8, and 9 present the original work. The equipment and results described in Chapter 9 are due partly to myself and partly to my predecessor Dr A.D. Seagar [Seagar 1983a]. His and my respective contributions are identified in the introduction to Chapter 9.

Part 1 of the thesis begins with Chapter 1, a review of fundamental theory pertaining to imaging systems in general. The general imaging system is subdivided into source, propagation medium, and imaging instrument, and these components are examined in turn. The material presented here underlies the whole of the first part of the thesis.

Chapter 2 discusses the imaging system of optical astronomy. Here the sources are celestial objects such as stars and planets, the propagation medium comprises outer space and (more importantly) the Earth's atmosphere, and the imaging instrument is an optical telescope. The effects of the atmosphere on astronomical images are described, and some classical and recent techniques for circumventing the "seeing problem" in high spatial resolution optical astronomy are reviewed.

One class of techniques for beating the seeing problem originated with Labeyrie's introduction of speckle interferometry in 1970. These astronomical speckle processing techniques are reviewed in detail in Chapter 3. This review is a broader, though necessarily less deep, compendium of astronomical speckle processing methods than any reviews I am aware of in the literature.

Many of the original results presented in Part 1 are founded on digital computer simulations of the optical astronomical imaging system. The simulations are concerned primarily with the optical properties of the Earth's atmosphere. Chapter 4 describes the algorithms and the model of the atmosphere employed in the simulations.

Chapter 5 presents new results on the behaviour of the shift-and-add speckle imaging technique of Bates and Cady [1980] under nonisoplanatic and photon-limited conditions. Each of these two deviations from ideal conditions is of significant importance in astronomical practice. The results are drawn from computer simulations and mathematical modelling, and from speckle processing of astronomical speckle images sent to this University by E.K. Hege and J.M. Beckers of the Steward Observatory in Tucson, Arizona.

In Chapter 6 a new speckle imaging technique called zero-and-add is described. Zero-and-add is one of the recent products of the continuing interest of Professor Bates and a number of his research students in the complex zeros of entire functions. Special attention is paid to the performance of zero-and-add when speckle images deviate in various ways from perfect isoplanatism.

Part 2 of the thesis, dealing with the medical side of my work, begins with Chapter 7. This chapter reviews techniques for measuring gas volume within, and gas flow into and out of, a person's lungs. Non-invasive transducers which respond to the movements of the chest and abdomen walls during breathing are given particular attention in the review. These transducers are suitable for long-term respiratory monitoring in clinical applications, such as the application described in Chapter 8.

Chapter 8 details the major part of the work which I carried out with the Christchurch Clinical School of Medicine. It describes a computerised system for long-term respiratory monitoring developed by myself within the Division of Anaesthesia of the Clinical School. The immediate purpose of this equipment is to allow a clinical study to be made of the respiratory effects of various forms of pain relief for hospital patients undergoing surgery. The system is potentially broadly applicable in clinical and laboratory respiratory studies. This chapter describes the system's non-invasive transducer of breathing motions, its microcomputer, its data processing algorithms, and preliminary clinical results obtained from the system during its development.

In Chapter 9 a microprocessor-controlled venous occlusion plethysmograph jointly developed by A.D. Seagar and myself is described. The electronic hardware and the data analysis software embodied in this instrument were developed by Seagar [1983a] during the part of his Ph.D. research undertaken with the Division of Anaesthesia of the Christchurch Clinical School of Medicine [cf. Seagar et al 1984]. On completing his Ph.D., Seagar began incorporating the hardware and software of the system he had developed into an automated, microprocessor-controlled system which would perform venous occlusion plethysmographic measurements and data analysis at the press of a button. The new, automated system was incomplete when Seagar left New Zealand to take up a position with the Royal Hallamshire Hospital in Sheffield, England. My contribution was to complete the construction, programming, and commissioning of the automated system after Seagar's departure. The system is now essentially complete and has been applied in both medical research and patient care in Christchurch hospitals. Chapter 9 describes the

finished system and presents examples of clinical results obtained with it.

Part 3 of the thesis, comprising Chapter 10, summarises the various conclusions which are drawn in detail in the individual chapters of Parts 1 and 2, and presents suggestions for further research.

The computer software which generated and displayed the results presented in this thesis was written entirely by myself, except for the shared contribution with Seagar to the firmware of the automated venous occlusion plethysmograph described in Chapter 9. Graphics display was performed through graphics device drivers written by Seagar (pen plotters) and D.G. Bailey (video display).

While the work reported in this thesis was in progress the following papers and presentations were prepared:

A.M. SINTON, P.H. GARDENIER, R.H.T. BATES
"Reinvestigation of optical interference at low light levels"
Speculations in Science and Technology, in press.

A.M. SINTON, J.M. GIBBS
"Breath by breath monitoring of respiratory patterns before and after surgery"
Presented to the Christchurch Medical Research Society, Christchurch, New Zealand, 17 July 1985. Abstract in The New Zealand Medical Journal, Vol. 98 No. 792, 11 December 1985. Pp. 1052 - 1053

R.H.T. BATES, A.M. SINTON, R.A. MINARD
"Generalization of shift-and-add imaging"
Proceedings of the SPIE, Vol. 556, 20 - 23 August 1985. Pp. 263 - 269

J.M. GIBBS, A.M. SINTON
"Inductance plethysmographic respiratory measurements in the postoperative period: methods and preliminary results"
Presented to the Annual Meeting of the Combined Anaesthetists in New Zealand, Auckland, New Zealand, November 1985.

A.M. SINTON, B.L.K. DAVEY, R.H.T. BATES
"Augmenting shift-and-add with zero-and-add"
Journal of the Optical Society of America A, to appear June 1986.

B.L.K. DAVEY, A.M. SINTON, R.H.T. BATES
"Zero-and-add"
Optical Engineering, in press.

A.M. SINTON, J.M. GIBBS
"Long term respiratory monitoring before and after anaesthesia"
Accepted for presentation (by J.M. Gibbs) at the 7th Asian Australasian Congress of Anaesthesiologists, Hong Kong, September 1986.

MATHEMATICAL GLOSSARY

This glossary defines certain mathematical operators and functions which are commonly used in image processing. Most of these definitions are invoked in the body of the thesis.

FOURIER TRANSFORM [cf. Goodman 1968 p.5; Bracewell 1978 p.251, cf. p.7]

$$\mathbf{F}[f(\underset{\sim}{x})] = \int_{-\infty}^{\infty} f(\underset{\sim}{x})\, e^{-j2\pi \underset{\sim}{u}.\underset{\sim}{x}} .d\underset{\sim}{x}$$

$$= F(\underset{\sim}{u})$$

$$\mathbf{F}^{-1}[F(\underset{\sim}{u})] = \int_{-\infty}^{\infty} F(\underset{\sim}{u})\, e^{j2\pi \underset{\sim}{u}.\underset{\sim}{x}} .d\underset{\sim}{u}$$

$$= f(x)$$

$$f(\underset{\sim}{x}) \leftrightarrow F(\underset{\sim}{u})$$

FOURIER-BESSEL TRANSFORM [Goodman 1968 p.13] or HANKEL TRANSFORM of zero order [Bracewell 1978 p.248]

$$\mathbf{B}[f(r)] = 2\pi \int_{0}^{\infty} f(r)\, J_0(2\pi\rho r) .dr$$

$$= F(\rho)$$

$$\mathbf{B}^{-1}[F(\rho)] = \mathbf{B}[F(\rho)] = 2\pi \int_{0}^{\infty} F(\rho)\, J_0(2\pi\rho r) .d\rho$$

$$= f(r)$$

CONVOLUTION [cf. Bracewell 1978 p.25]

$$f(\underset{\sim}{x}) \circledcirc g(\underset{\sim}{x}) = \int_{-\infty}^{\infty} f(\underset{\sim}{x}')\, g(\underset{\sim}{x} - \underset{\sim}{x}') .d\underset{\sim}{x}'$$

$$= g(\underset{\sim}{x}) \circledcirc f(\underset{\sim}{x})$$

$$\leftrightarrow F(\underset{\sim}{u})\, G(\underset{\sim}{u})$$

CORRELATION [cf. Bracewell 1978 p.46]

$$fg(\underset{\sim}{x}) = f(\underset{\sim}{x}) \otimes g(\underset{\sim}{x}) = \int_{-\infty}^{\infty} f^*(\underset{\sim}{x}') g(\underset{\sim}{x} + \underset{\sim}{x}') \, .d\underset{\sim}{x}'$$

$$= \int_{-\infty}^{\infty} f^*(\underset{\sim}{x}'' - \underset{\sim}{x}) g(\underset{\sim}{x}'') \, .d\underset{\sim}{x}''$$

$$= f^*(-\underset{\sim}{x}) \odot g(\underset{\sim}{x})$$

$$= (gf(-\underset{\sim}{x}))^*$$

$$\leftrightarrow F^*(\underset{\sim}{u}) G(\underset{\sim}{u})$$

AUTOCORRELATION [cf. Bracewell 1978 p.41]

$$\mathbf{A}[f(\underset{\sim}{x})] = ff(\underset{\sim}{x}) = f(\underset{\sim}{x}) \otimes f(\underset{\sim}{x}) = \int_{-\infty}^{\infty} f^*(\underset{\sim}{x}') f(\underset{\sim}{x} + \underset{\sim}{x}') \, .d\underset{\sim}{x}'$$

$$\leftrightarrow F^*(\underset{\sim}{u}) F(\underset{\sim}{u}) = |F(\underset{\sim}{u})|^2$$

NORMALISED AUTOCORRELATION [cf. Bracewell 1978 p.41]

$$\hat{\mathbf{A}}[f(\underset{\sim}{x})] = \frac{ff(\underset{\sim}{x})}{ff(\underset{\sim}{0})} = \frac{\int_{-\infty}^{\infty} f^*(\underset{\sim}{x}) f(\underset{\sim}{x}' + \underset{\sim}{x}') \, .d\underset{\sim}{x}'}{\int_{-\infty}^{\infty} f^*(\underset{\sim}{x}') f(\underset{\sim}{x}') \, .d\underset{\sim}{x}'}$$

RECTANGLE FUNCTION [Goodman 1968 p.13, cf. Bracewell 1978 p.52]

$$rect(x) = \begin{cases} 1 & |x| \leq \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}$$

TRIANGLE FUNCTION [Goodman 1968 p.14, cf. Bracewell 1978 p.53]

$$\Lambda(x) = \begin{cases} 1 - |x| & |x| \leq 1 \\ 0 & \text{otherwise} \end{cases} = \mathbf{A}[rect(x)]$$

SINC FUNCTION [Goodman 1968 p.14, Bracewell 1978 p.62]

$$sinc(x) = \frac{\sin(\pi x)}{\pi x} = \mathbf{F}[rect(x)]$$

CIRCLE FUNCTION [cf. Goodman 1968 p.14]

$$\mathrm{circ}(x,y) = \mathrm{rect}(r); \quad r = \sqrt{x^2 + y^2}$$

JINC FUNCTION [Bracewell 1978 p.421]

$$\mathrm{jinc}(r) = \frac{J_1(\pi r)}{2r} = \mathbf{B}[\mathrm{rect}(\rho)] = \mathbf{F}[\mathrm{circ}(u,v)]$$

MEDICAL GLOSSARY

This glossary gives definitions of common medical terms that I have encountered during the research associated with Chapters 7, 8, and 9 of this thesis. It is primarily intended as an aid to reading the medical papers quoted in the references of this thesis. It may also be of use when reading Chapters 7, 8, and 9 themselves, in which medical terms occasionally appear without prior definition where knowledge of their meanings is not essential to the narrative. Most of the definitions have been taken straight from a medical dictionary [Friel 1974], but some of the definitions have been modified or concocted to suit the special meanings of the terms in the contexts of Chapters 7 to 9.

ACETABULUM : cup-shaped cavity in the hip bone, into which the head of the femur fits forming the hip joint

AIRWAY : the route for passage of air into and out of the lungs

ANAESTHESIA : pain relief, including pain relief with other effects such as loss of consciousness and muscle relaxation

ANAESTHETIC : agent (especially a drug) which causes anaesthesia

ANALGESIA : pain relief, especially pain relief without loss of consciousness

ANALGESIC : agent (especially a drug) which causes analgesia

ANTERIOR : front (cf. posterior)

APNOEA : absence of ventilation of the lungs

ARTEFACT : detail or feature in a physiological signal which does not come from the expected source

BASAL (SIGNAL) : a physiological signal from a transducer whose frequency response extends down to DC and thus includes the DC or "baseline" component

BREATHING : the inflation and deflation of the lungs

BRONCHI : plural of bronchus

BRONCHOSPASM : spasm of the smooth muscles of the bronchi, as occurs in asthma

BRONCHUS : (Greek "windpipe") any of the larger air passages of the lungs

CENTRAL APNOEA : apnoea accompanied by a cessation of breathing motions (cf. obstructive apnoea)

CENTRAL NERVOUS SYSTEM : the brain and spinal cord

CHOLECYSTECTOMY : surgical removal of the gall bladder

CIRCULATION : (usually) movement of the blood through the heart and blood vessels

CIRCULATORY SYSTEM : (usually) the heart and blood vessels

CLINICAL : pertaining to the bedside, i.e. to the observation and treatment of patients rather than to theoretical or basic sciences

DISTAL : further from a reference point (cf. proximal)

DVT : Deep Vein Thrombosis

EMBOLISM : sudden blockage of a blood vessel by an embolus

EMBOLUS : blood clot or other foreign body floating in the blood

END-TIDAL CO_2 : carbon dioxide concentration of the gas from the lungs at the end of expiration, from which the carbon dioxide content of the arterial blood may be estimated

EPIDURAL (INJECTION) : injection at the spinal cord, outside its sheath (the dura mater) (cf. intrathecal)

FEMUR : long bone in the thigh, extending from the hip to the knee

GASTRECTOMY : surgical removal of part or all of the stomach

GASTRIC : pertaining to the stomach

HAEMOGLOBIN : chemical compound in blood which combines reversibly with oxygen and provides most of the oxygen carrying capacity of the blood

HYPERCAPNIA : excess of carbon dioxide in the blood

HYPOXAEMIA : deficiency of oxygen in the blood

HYPOXIA : deficiency of oxygen (e.g. in the inspired air, in the blood)

INTRATHECAL (INJECTION) : injection into the spinal cord, inside its sheath (the dura mater) (cf. epidural)

INVASIVE : (strictly) involving insertion of an instrument or foreign material into the body; (loosely) causing appreciable inconvenience or discomfort

IPG : Impedance Plethysmography

ISCHAEMIA : deficiency of blood in a part of the body due to constriction or obstruction of a blood vessel

ISCHAEMIC : pertaining to ischaemia

LATERAL : side or sideways

MINUTE VENTILATION (or MINUTE VOLUME) : total volume of gas inspired or expired in one minute

MORBID : pertaining to disease; diseased

MORBIDITY : condition of being diseased

MORBIDITY (RATE) : proportion of sick persons in a group

MORTALITY (RATE) : proportion of persons in a group who die in a given interval

NARCOTIC : drug which causes narcosis, i.e. stupor or loss of sensibility

OBSTRUCTED MOTIONS : abnormal breathing motions which occur when the airways to the lungs are partly or completely blocked; usually includes some degree of paradoxical motion

OBSTRUCTIVE APNOEA : apnoea accompanied by obstructed breathing motions (cf. central apnoea)

OPIOID : drug with morphine-like properties, not necessarily derived from opium (i.e. natural or synthetic)

OPIATE : drug with morphine-like properties, derived from opium (i.e. natural)

OPIUM : drug obtained from the juice of the opium poppy, containing morphine and other addictive compounds

ORAL : by mouth (cf. parenteral)

PAIN : sensation of discomfort, distress, or agony resulting from the stimulation of specialised nerve endings

PARADOXICAL MOTION : abnormal breathing pattern in which the chest and abdomen walls move in opposite directions, instead of expanding and contracting together as is normal

PARENTERAL : entering the body by a route other than the alimentary canal, i.e. by injection (cf. oral)

PLETHYSMOGRAPH : device for measuring volume changes in the body

PNEUMOTACHOGRAPH : device for measuring gas flow to and from the lungs

POSTERIOR : back (cf. anterior)

PROGNOSIS : forecast of outcome, or prospect of recovery, from a disease or injury

PRONE : lying face downward (cf. supine)

PROSTHESIS : an artificial substitute for a body part

PROXIMAL : nearer to some reference point (cf. distal); e.g. proximal deep vein thrombosis, occurring closer to the heart than a measurement site in the calf

PULMONARY : pertaining to the lungs

RESPIRATION : the exchange of oxygen and carbon dioxide between the atmosphere and the cells of the body

RESPIRATORY DEPRESSION : changes in respiration, induced for example by anaesthetic drugs, which potentially or actually impair the efficiency of respiration

SPASM : sudden, involuntary muscle contraction

SPIROMETER : a device for measuring changes in lung volume: an extensible chamber into which the subject breathes

SUPINE : lying face upward (cf. prone)

SURGERY : the treatment of diseases, injuries, and deformities by manual or operative means

SYSTEMIC : affecting the body as a whole

THROMBOSIS : the formation, development, or presence of a thrombus

THROMBUS : (loosely) a blood clot

TIBB : Time Interval Between Breaths: the period, or duration, of a breath

TIDAL VOLUME : increase (or decrease) of lung volume associated with the inspiration (or expiration) of a single breath

TRAUMA : any physical or psychic wound or injury

VASCULAR : pertaining to blood vessels

VASCULATURE : the vascular system of the body or any part of it

VASOACTIVE : exerting an effect on the diameter of blood vessels

VENOUS : pertaining to the veins

VENTILATION : the exchange of gas, or (quantitatively) the volume or flow of gas exchanged, between the lungs and the atmosphere

VOP : Venous Occlusion Plethysmography, a non-invasive technique for measuring circulation in the limbs

Part 1

IN ASTRONOMY

CHAPTER 1

THE GENERAL IMAGING SYSTEM

1.1 INTRODUCTION

This chapter presents a review of fundamental principles relating to imaging systems. The results presented here form a foundation for the discussions of the optical astronomical imaging system in Chapters 2 to 6, but also apply generally to any imaging system in which the radiation from the source is appropriately described by a linear, scalar wave equation.

Section 1.2 presents briefly the concepts of linear systems and the analytic signal, which are central to what follows. The component parts of the general imaging system are introduced in Section 1.3. The source, the propagation medium, and the imaging instrument are discussed in Sections 1.4, 1.5, and 1.6 respectively.

1.2 PRELIMINARIES

1.2.1 LINEAR SYSTEMS

A system possessing an input $f(\underset{\sim}{x})$ and an output $s(\underset{\sim}{\alpha})$ is described by the transformation **S** which maps the input to the output. The mapping process is written as

$$s(\underset{\sim}{\alpha}) = \mathbf{S}[f(\underset{\sim}{x})]. \tag{1.1}$$

$\underset{\sim}{x}$ and $\underset{\sim}{\alpha}$ are representative members of different sets of independent variables, and are of arbitrary and not necessarily equal dimensionality. The system is assumed to produce a unique output for each input, i.e. the transformation **S** is deterministic (non-random) and one-to-one or many-to-one from f to s.

Linear systems are those for which the relation

$$S[af_1(\underset{\sim}{x})+bf_2(\underset{\sim}{x})] = aS[f_1(\underset{\sim}{x})] + bS[f_2(\underset{\sim}{x})] \qquad (1.2)$$

is true for all inputs f_1 and f_2 and for all constants a and b. The response of a linear system to a weighted sum of inputs is equal to the weighted sum of its responses to each input applied separately.

Linear systems are important because the mathematical tools for handling them are well developed. A useful tool is the superposition integral, in which the input is decomposed into a collection of elementary inputs and the output is expressed as the sum of the outputs to each elementary input considered in isolation. Appropriate choice of elementary inputs simplifies the solution of problems concerning the system's behaviour.

Frequently, an appropriate choice of elementary inputs is based on the Dirac delta function $\delta(\underset{\sim}{x})$ [Bracewell 1978 ∅5], which is a function having effectively nonzero value only for one particular value of $\underset{\sim}{x}$. $\delta(\underset{\sim}{x})$ is defined here by the integrals

$$\int_{T+} \delta(\underset{\sim}{x}).d\underset{\sim}{x} = 1$$

$$\int_{T-} \delta(\underset{\sim}{x}).d\underset{\sim}{x} = 0 \qquad (1.3)$$

where T^+ is an infinitesimally small region of $\underset{\sim}{x}$-space containing $\underset{\sim}{x} = \underset{\sim}{0}$, and T^- is any region not containing T^+. $\delta(\underset{\sim}{x})$ possesses the property

$$\int_T f(\underset{\sim}{x})\delta(\underset{\sim}{x}-\underset{\sim}{x}').d\underset{\sim}{x} = f(\underset{\sim}{x}') \quad \forall\, T : \underset{\sim}{x}' \in T \qquad (1.4)$$

which is known as the sifting property. In the superposition integral, the elementary components of the input are chosen to be the weighted, translated delta functions $f(\underset{\sim}{x}')\delta(\underset{\sim}{x}-\underset{\sim}{x}')$. The input as a whole is equal to the sum of all of these, i.e.

$$f(\underset{\sim}{x}) = \int f(\underset{\sim}{x}')\delta(\underset{\sim}{x}-\underset{\sim}{x}').d\underset{\sim}{x}', \qquad (1.5)$$

which is just a reinterpretation of the sifting property (1.4). The response of the system to a single translated delta function is called the point spread function or psf, defined by

$$h(\underset{\sim}{\alpha},\underset{\sim}{x}') = \mathbf{S}[\delta(\underset{\sim}{x}-\underset{\sim}{x}')]. \tag{1.6}$$

The output as a whole is given by

$$\begin{aligned} s(\underset{\sim}{\alpha}) &= \mathbf{S}[f(\underset{\sim}{x})] \\ &= \mathbf{S}\left[\int f(\underset{\sim}{x}')\delta(\underset{\sim}{x}-\underset{\sim}{x}').d\underset{\sim}{x}'\right] \\ &= \int \mathbf{S}[f(\underset{\sim}{x}')\delta(\underset{\sim}{x}-\underset{\sim}{x}')].d\underset{\sim}{x}' \qquad (1.7a) \\ &= \int f(\underset{\sim}{x}')\mathbf{S}[\delta(\underset{\sim}{x}-\underset{\sim}{x}')].d\underset{\sim}{x}' \qquad (1.7b) \\ &= \int f(\underset{\sim}{x}')h(\underset{\sim}{\alpha},\underset{\sim}{x}').d\underset{\sim}{x}' \qquad (1.8) \end{aligned}$$

where the linearity of **S** has been invoked in (1.7a) and (1.7b). Equation (1.8) is a superposition integral.

An important subset of the linear systems are those for which $\underset{\sim}{\alpha}$ and $\underset{\sim}{x}$ have the same dimensionality and for which the value of the point spread function $h(\underset{\sim}{\alpha},\underset{\sim}{x}')$ depends only in the difference between $\underset{\sim}{\alpha}$ and $\underset{\sim}{x}'$, i.e.

$$h(\underset{\sim}{\alpha},\underset{\sim}{x}') = h(\underset{\sim}{\alpha}-\underset{\sim}{x}'). \tag{1.9}$$

Such systems are termed point spread invariant (psi) or isoplanatic [Welford 1976], because the shape of the point spread function in $\underset{\sim}{\alpha}$-space is the same for all values of $\underset{\sim}{x}'$. The converse, general condition expressed by (1.8) is termed point spread variant (psv) or nonisoplanatic. Often the shape of the point spread function changes only slowly with $\underset{\sim}{x}'$, in which case the system can be considered to be approximately isoplanatic over limited regions called isoplanatic patches. Substituting (1.9) into (1.8) gives, for an isoplanatic or psi system,

$$s(\underset{\sim}{x}) = \int f(\underset{\sim}{x}')h(\underset{\sim}{x}-\underset{\sim}{x}').d\underset{\sim}{x}' \tag{1.10}$$

where $\underset{\sim}{x}$ has been written in place of $\underset{\sim}{\alpha}$ because the two now have the same dimensionality and the distinction between them can be dropped. Equation (1.10) is a convolution integral, for which there is the explicit notation

$$s(x) = f(x) \odot h(x), \tag{1.11}$$

with the symbol $\odot$ meaning convolution.

A powerful tool in the treatment of point-spread invariant systems is the Fourier transform [Bracewell 1978], here denoted by the boldface **F** and defined as

$$\text{FORWARD TRANSFORM} \quad \mathbf{F}[f(\underset{\sim}{x})] = \int f(\underset{\sim}{x}) e^{-j2\pi \underset{\sim}{u}\cdot\underset{\sim}{x}} \, d\underset{\sim}{x} = F(\underset{\sim}{u}), \tag{1.12a}$$

$$\text{REVERSE TRANSFORM} \quad \mathbf{F}^{-1}[F(\underset{\sim}{u})] = \int F(\underset{\sim}{u}) e^{j2\pi \underset{\sim}{u}\cdot\underset{\sim}{x}} \, d\underset{\sim}{u} = f(\underset{\sim}{x}). \tag{1.12b}$$

The dimensionalities of $\underset{\sim}{x}$ and $\underset{\sim}{u}$ are the same. The Fourier transform is itself a linear operator since

$$\mathbf{F}[af_1(\underset{\sim}{x}) + bf_2(\underset{\sim}{x})] = a\mathbf{F}[f_1(\underset{\sim}{x})] + b\mathbf{F}[f_2(\underset{\sim}{x})] \tag{1.13}$$

(cf. 1.2). The common convention of denoting functions which are Fourier transforms of each other by different cases of the same letter, and the Fourier transform relationship between them by a double-ended arrow, is followed in this thesis, e.g. $f(\underset{\sim}{x}) \leftrightarrow F(\underset{\sim}{u})$ are a Fourier transform pair.

If the Fourier transform is applied to the left and right hand sides of (1.11), the convolution theorem [Bracewell 1978 pp.108-] states that the result is

$$S(\underset{\sim}{u}) = F(\underset{\sim}{u})\, H(\underset{\sim}{u}) \tag{1.14}$$

where $f(\underset{\sim}{x}) \leftrightarrow F(\underset{\sim}{u})$, $h(\underset{\sim}{x}) \leftrightarrow H(\underset{\sim}{u})$, and $s(\underset{\sim}{x}) \leftrightarrow S(\underset{\sim}{u})$. The function $H(\underset{\sim}{u})$, the Fourier transform of the point spread function, is called the transfer function of the system. $F(\underset{\sim}{u})$ and $S(\underset{\sim}{u})$ are the input and output spectra. In an imaging context, F and S are sometimes referred to as the input and output visibilities because in many situations it

is the Fourier domain quantities that are directly measurable or "visible". By means of the Fourier transform, the relatively complicated convolution expression (1.10) for the system output has been reduced to a product (1.14). The product form is easier to handle in algebraic manipulation, and can be much quicker to compute in numerical evaluation with the aid of the Fast Fourier Transform (FFT) [Brigham 1974].

1.2.2 THE ANALYTIC SIGNAL

Signals measured in the "real world" are real functions of space and time, for example the output voltage $v_r(\underset{\sim}{x},t)$ at time t of a transducer placed at position $\underset{\sim}{x}$. A signal having a single temporal frequency ν, i.e. a monochromatic signal, is represented by the sinusoidal function

$$v_r(\underset{\sim}{x},t) = A(\underset{\sim}{x}) \cos(2\pi\nu t+\phi(\underset{\sim}{x})) \tag{1.15}$$

with $A(\underset{\sim}{x})$ being the (real) amplitude of the signal and $\phi(\underset{\sim}{x})$ being its (real) phase. It is often convenient to consider $v_r(\underset{\sim}{x},t)$ as the real part of a complex exponential function,

$$v_r(\underset{\sim}{x},t) = \mathrm{Re}[v(\underset{\sim}{x},t)] :$$

$$v(\underset{\sim}{x},t) = A(\underset{\sim}{x})\, e^{j(2\pi\nu t+\phi(\underset{\sim}{x}))} , \tag{1.16}$$

because the complex exponential function is mathematically easy to handle. When solving problems involving linear operations on v_r it is possible to perform the same operations on v, which may be easier than on v_r, to obtain a complex result whose real part yields the desired real result. Thus the real signal is well represented by $v(\underset{\sim}{x},t)$.

The monochromatic complex signal (1.16) may be rewritten as

$$v(\underset{\sim}{x},t) = A(\underset{\sim}{x})\, e^{j\phi(\underset{\sim}{x})}\, e^{j2\pi\nu t} \tag{1.17}$$

in which form the time dependence appears only in the second

exponential term. When the time dependence is known (i.e. the frequency of the monochromatic signal is known or can be assumed), it can be omitted from the explicit notation - it is still present implicitly - by writing

$$\mathbf{v}(\underset{\sim}{x}) = A(\underset{\sim}{x})\, e^{j\phi(\underset{\sim}{x})} \tag{1.18}$$

where $\mathbf{v}(\underset{\sim}{x})$ is the phasor representing $v(\underset{\sim}{x},t)$.

The analytic signal provides a generalisation of the monochromatic complex-signal representation (1.17, 1.18) to polychromatic signals. Let $v_r(\underset{\sim}{x},t)$ be an arbitrary real function. The analytic signal associated with v_r is the complex function [Bracewell 1978 p.268]

$$v(\underset{\sim}{x},t) = v_r(\underset{\sim}{x},t) - jv_i(\underset{\sim}{x},t) \tag{1.19}$$

whose real part is v_r and whose imaginary part is obtained from the Hilbert transform of v_r, defined by

$$v_i(\underset{\sim}{x},t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{v_r(\underset{\sim}{x},t')}{t' - t} .dt' \tag{1.20}$$

$$= \frac{-1}{\pi t} \odot v_r(\underset{\sim}{x},t) \tag{1.21}$$

(the convolution is in the time domain). The Cauchy principal value of the integral is taken to account for the singularity at $t = t'$. v_i is the quadrature function or conjugate function of v_r. The analytic signal is so called because when treated as a function of a complex variable, by making t complex-valued in (1.19), it is analytic in half of the complex plane. It contains no negative temporal frequencies, i.e.

$$V(\underset{\sim}{x},\nu) = \int_{-\infty}^{\infty} v(\underset{\sim}{x},t)\, e^{-j2\pi\nu t} .dt = 0 \quad \forall\ \nu < 0 \tag{1.22}$$

A quasi-monochromatic signal is one whose spectral components

have appreciable value only over a frequency interval $\Delta\nu$ much smaller than the centre frequency $\bar{\nu}$. Its analytic signal can be usefully written in the form [cf. Born and Wolf 1975 φ10.2]

$$v(\underset{\sim}{x},t) = A(\underset{\sim}{x},t)\, e^{j\phi(\underset{\sim}{x},t)}\, e^{j2\pi\bar{\nu}t} , \tag{1.23}$$

where $A(\underset{\sim}{x},t)$ represents the envelope or time-varying amplitude of the real signal and $\phi(\underset{\sim}{x},t)$ represents its time-varying phase. As for the monochromatic signal (1.17), when the mean frequency $\bar{\nu}$ can be assumed the quasi-monochromatic signal can be represented by a slowly time-varying generalised phasor

$$\mathbf{v}(\underset{\sim}{x},t) = A(\underset{\sim}{x},t)\, e^{j\phi(\underset{\sim}{x},t)} \tag{1.24}$$

with the $e^{j2\pi\bar{\nu}t}$ understood.

The analytic signal and the properties of the Hilbert transform are discussed informatively by Beran and Parrent [1974 φφ2.1-2.3]. Bracewell [1978 pp.267-] illustrates the use of the Hilbert transform to define the instantaneous frequency of a polychromatic signal.

1.3 COMPONENTS OF THE GENERAL IMAGING SYSTEM

The components of a general imaging system are shown in Figure 1.1. The source is a material body or physical process from which comes radiation, or emanations. The radiation propagates from the source through a medium. Some of the radiation reaches the imaging instrument, and is received through the imaging instrument's aperture.

The object, which is the quantity the imaging system is designed to reveal information about, is some physical property of the source that is of interest to the observer, for example the spatial distribution of intensity of radiation from the source. The image, which is formed by the imaging system from the radiation incident upon its aperture, is a measurable physical quantity containing information about the object.

A physical picture of an imaging system is shown in Figure 1.2. The medium is shown in a localised region for clarity, but in fact extends from the source to the imaging instrument. In this picture the source and the medium occupy three-dimensional space but the image formed is two-dimensional, as is commonly the case in practice (e.g. when imaging with a conventional astronomical telescope). The aperture of the imaging instrument is considered to lie in a surface called the aperture plane, and the image is formed in the image plane. A two-dimensional object lying in the object plane is defined by means of the equivalent source described in Section 1.4.1.

1.4 THE SOURCE

Physically, the source in a general imaging system (Figure 1.1) is some "thing" producing radiation which propagates through space away from the source. The source may be self-luminous, with the energy that generates the radiation coming from the source itself, or it may be a passive scatterer of radiation impinging upon it from an external generator. Mathematically, the source is represented by a source density Σ describing the emission from each part of the source at each instant.

1.4.1 THE EQUIVALENT SOURCE

In many imaging systems, including the majority of the situations in optical astronomy described in this thesis, the image is

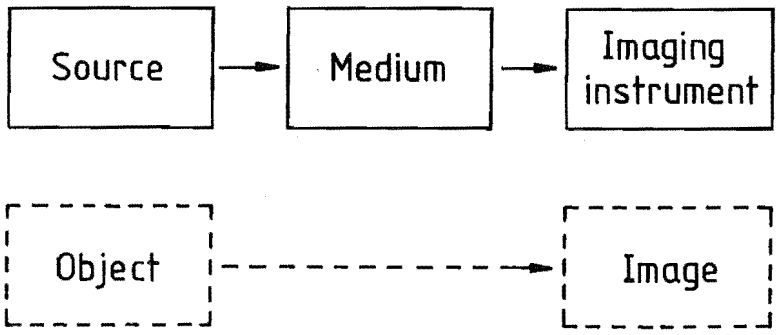


Figure 1.1: Conceptual picture of a general imaging system.

formed on a two-dimensional surface. A physical picture of such imaging systems is provided in Figure 1.2. The source density ${}^3\Sigma({}^3\underset{\sim}{x},t)$ (Figure 1.2) is a function of time and three-dimensional space. The leading superscript numbers are used to denote the dimensionality of a quantity, either a position vector or a function of a position vector. Thus ${}^3\underset{\sim}{x}$ is a three-dimensional position vector, ${}^3\underset{\sim}{x} = (x,y,z)$, and ${}^3\Sigma$ is a function of three-dimensional position. This notation is used below to distinguish between two- and three-dimensional quantities.

For imaging systems forming a two-dimensional image it is useful to define a two-dimensional equivalent source with density ${}^2\Sigma({}^2\underset{\sim}{x},t)$ (Figure 1.2) which produces the same effect upon the imaging instrument that ${}^3\Sigma({}^3\underset{\sim}{x},t)$ does. A two-dimensional object can then be associated with the properties of ${}^2\Sigma$. Huygen's principle as expressed through Green's theorems [Silver 1965 ∅4.1, Born and Wolf 1975 ∅8.3] confirms that a ${}^2\Sigma({}^2\underset{\sim}{x},t)$ can be chosen, lying on a surface Π that separates space into two halves with the source ${}^3\Sigma({}^3\underset{\sim}{x},t)$ entirely in one half and the rest of the imaging system entirely in the other half, whose radiation is identical with that from ${}^3\Sigma({}^3\underset{\sim}{x},t)$ on the

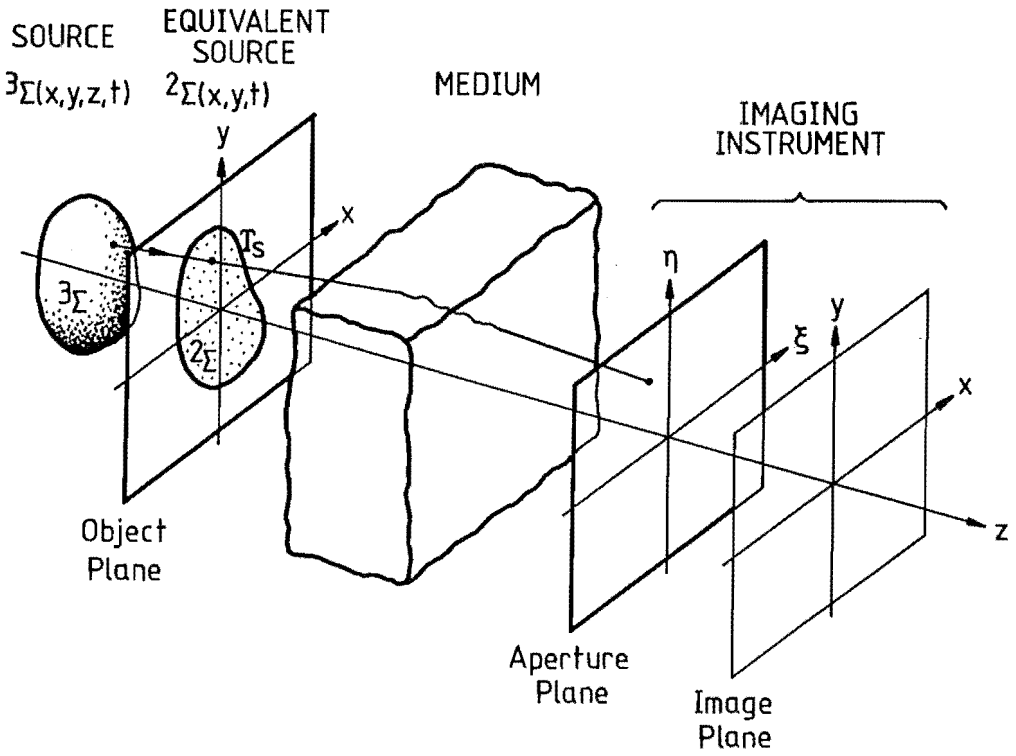


Figure 1.2: Physical picture of a general imaging system.

imaging instrument's side of Π (Figure 1.2). $^2\underset{\sim}{x}$ is a two-dimensional position vector specifying position on Π. The surface Π may be thought of as enclosing the source (e.g. in microscopy), or as enclosing the imaging instrument (e.g. the celestial sphere of astronomy), or as an infinite plane as in Figure 1.2. When Π is planar, or can be considered so, for example when viewing a small portion of the celestial sphere, it is appropriately called the object plane. ^2T_S is the region on Π outside which $^2\Sigma(^2\underset{\sim}{x},t)$ is zero. In many situations ^2T_S can be considered to be of finite size, as for example in the aperture-field method [Silver 1965 ϕ5.11] where, as a high-frequency approximation, ^2T_S is taken to be the geometrical projection onto Π of the extent ^3T_S of a source lying just behind Π, formed by all the rays that reach the imaging instrument from the source.

Since the effects of the (true) source $^3\Sigma$ and the equivalent source $^2\Sigma$ on the imaging instrument are identical by definition, from now on the term "source" is used to refer to both $^3\Sigma$ and $^2\Sigma$. When the unsubscripted notation $\Sigma(\underset{\sim}{x},t)$ is used it may refer to either $^2\Sigma$ or $^3\Sigma$.

1.4.2 COHERENCE

The concept of coherence describes the synchronism between radiation ψ at different points in space and time. The complex mutual coherence function can be defined as [Beran and Parrent 1974 ϕ2.5]

$$\Gamma(\underset{\sim}{x}_1,\underset{\sim}{x}_2,\Delta t) = \Gamma_{12}(t) = \langle\psi^*(\underset{\sim}{x}_1,t)\psi(\underset{\sim}{x}_2,t+\Delta t)\rangle \qquad (1.25)$$

The concept may be applied directly to the source density,

$$\Gamma(\underset{\sim}{x}_1,\underset{\sim}{x}_2,\Delta t) = \langle\Sigma^*(\underset{\sim}{x}_1,t)\Sigma(\underset{\sim}{x}_2,t+\Delta t)\rangle \qquad (1.26)$$

The complex degree of coherence, or normalised mutual coherence function, is defined as

$$\gamma(\underset{\sim}{x}_1,\underset{\sim}{x}_2,\Delta t) = \gamma_{12}(\Delta t) = \frac{\Gamma_{12}(\Delta t)}{\sqrt{\Gamma_{11}(0)\Gamma_{22}(0)}} . \qquad (1.27)$$

The averages < > in (1.25) and (1.26) are time averages taken over a sufficiently long interval that random fluctuations in the averaged quantity become negligible. This formulation carries the implicit assumption that the statistics of $\Sigma(\underset{\sim}{x},t)$ are stationary with time.

Coherence over time intervals and over space intervals may be considered separately as temporal coherence and spatial coherence respectively. The temporal coherence or self coherence at a point is defined by

$$\Gamma_T(\underset{\sim}{x},\Delta t) = \Gamma(\underset{\sim}{x},\underset{\sim}{x},\Delta t) = \langle\psi^*(\underset{\sim}{x},t)\psi(\underset{\sim}{x},t+\Delta t)\rangle, \tag{1.28a}$$

$$\gamma_T(\underset{\sim}{x},\Delta t) = \frac{\Gamma_T(\underset{\sim}{x},\Delta t)}{\Gamma_T(\underset{\sim}{x},0)} = \gamma(\underset{\sim}{x},\underset{\sim}{x},\Delta t). \tag{1.28b}$$

The spatial coherence, or mutual coherence, or mutual intensity between two points is defined by

$$\Gamma_S(\underset{\sim}{x}_1,\underset{\sim}{x}_2) = \Gamma(\underset{\sim}{x}_1,\underset{\sim}{x}_2,0) = \langle\psi^*(\underset{\sim}{x}_1,t)\psi(\underset{\sim}{x}_2,t)\rangle, \tag{1.29a}$$

$$\gamma_S(\underset{\sim}{x}_1,\underset{\sim}{x}_2) = \gamma(\underset{\sim}{x}_1,\underset{\sim}{x}_2,0). \tag{1.29b}$$

1.4.2.1 Temporal Coherence, Monochromaticity

The temporal coherence function $\gamma_T(\underset{\sim}{x},\Delta t)$ is equal to unity for $\Delta t = 0$ and, in general, $|\gamma_T|$ takes on values between 0 and 1 for $\Delta t > 0$. The effective width (with respect to Δt) of $\gamma_T(\underset{\sim}{x},\Delta t)$ is called the coherence time τ_c of the radiation or the source (see, for example, Born and Wolf [1975 ϕ7.5.8] for a discussion of interference experiments with spectral emission lines in which temporal coherence is measured, and Born and Wolf [1975 ϕ10.7.3] for a particular definition of coherence time). For convenience, τ_c is assumed to be the same for all points in the source. When τ_c is infinite, i.e.

$$|\gamma_T(\underset{\sim}{x},\Delta t)| = 1 \quad \forall\ \Delta t, \tag{1.30}$$

the source is monochromatic and the source density may be described by

$$\Sigma(\underset{\sim}{x},t) = A(\underset{\sim}{x})\, e^{j2\pi\nu t} \tag{1.31}$$

where ν is the temporal frequency of the radiation. $A(\underset{\sim}{x})$ is a complex scalar representing spatial variations in amplitude and phase within the source. When τ_c is finite but large compared to the mean period of the radiation from the source, the source is quasi-monochromatic and the source density may be represented by

$$\Sigma(\underset{\sim}{x},t) = A(\underset{\sim}{x},t)\, e^{j2\pi\bar{\nu}t} \tag{1.32}$$

where $\bar{\nu}$ is the mean temporal frequency of the radiation. In this generalisation, A has become a slowly varying function of time. The coherence time represents the shortest time interval over which the value of A can change appreciably.

Associated with the coherence time τ_c is a bandwidth $\Delta\nu$, whose value is of the order of the reciprocal of τ_c, depending on the particular definitions used for coherence time and bandwidth. One may write

$$\Delta\nu = \frac{1}{\tau_c} . \tag{1.33}$$

If the speed of propagation c of the radiation is introduced, length quantities related to the time quantities above may be defined:

coherence length $\quad L_c = c\,\tau_c$, (1.34)

mean wavelength $\quad \bar{\lambda} = \frac{c}{\bar{\nu}}$, (1.35)

bandwidth $\quad \Delta\lambda = \frac{c\Delta\nu}{\bar{\nu}^2} = \frac{\bar{\lambda}^2\Delta\nu}{c}$. (1.36)

The term "bandwidth" is used for both $\Delta\nu$ and $\Delta\lambda$; which one is meant can be determined from the context (time or length) in which the term is used.

The quasi-monochromatic condition is expressed as

$$\Delta\nu << \bar{\nu}, \tag{1.37a}$$

$$\Delta\lambda << \bar{\lambda}. \tag{1.37b}$$

The propagation of quasi-monochromatic radiation (1.32) may be treated using the same laws as apply to monochromatic radiation (1.31), with ν replaced by $\bar{\nu}$, provided that path length differences in the imaging system do not exceed the coherence length L_c [Beran and Parrent 1974 ϕ4.3].

1.4.2.2 Spatial Coherence

The spatial coherence function $\gamma_s(\underset{\sim}{x}_1, \underset{\sim}{x}_2)$ describes the degree of synchronism of the source or its radiation at different positions. $|\gamma_s|$ can in general take any value between 0 and 1. When applied to the source density Σ, two limiting cases are of particular interest.

When

$$|\gamma_s(\underset{\sim}{x}_1, \underset{\sim}{x}_2)| = 1 \quad \forall\ \underset{\sim}{x}_1, \underset{\sim}{x}_2 \in T_s, \tag{1.38}$$

the source is <u>spatially coherent</u>. The radiations from all points in the source occur in unison. As an example, any passive source reflecting spatially coherent radiation (e.g. radiation from another source at a single point) is itself spatially coherent, if its reflectivity is not varying randomly with time.

When

$$\gamma_s(\underset{\sim}{x}_1, \underset{\sim}{x}_2) = 0 \quad \forall\ |\underset{\sim}{x}_1 - \underset{\sim}{x}_2| > L;\ \underset{\sim}{x}_1, \underset{\sim}{x}_2 \in T_s \tag{1.39}$$

where L is a distance too small to be resolved (ϕ1.6.6) by the imaging system, the source is <u>spatially incoherent</u>. Radiations from discernibly different points are statistically independent of each other. Bodies emitting thermal radiation, such as incandescent lamps and stars, are spatially incoherent sources because the groups of atoms or molecules at resolvably separated points do indeed radiate

independently of each other. Spatially incoherent sources may be described by the idealised equation [cf. Beran and Parrent 1974 ϕ4.4]

$$\Gamma_S(\underset{\sim}{x}_1,\underset{\sim}{x}_2) = I(\underset{\sim}{x}_1)\delta(\underset{\sim}{x}_1-\underset{\sim}{x}_2) \qquad (1.40)$$

where

$$I(\underset{\sim}{x}_1) = \langle\psi^*(\underset{\sim}{x}_1,t)\psi(\underset{\sim}{x}_1,t)\rangle = \langle|\psi(\underset{\sim}{x}_1,t)|^2\rangle$$

is the intensity of the source or field at position $\underset{\sim}{x}_1$.

1.5 THE PROPAGATION MEDIUM

The source and the imaging instrument are imbedded in a propagation medium, which provides a vehicle for radiation to be carried from the source to the imaging instrument. The source perturbs the medium from its equilibrium state in the vicinity of the source, and these perturbations propagate away from the source through the medium as waves. The medium may be a material substance, such as a compressible fluid carrying pressure waves in acoustic imaging, or it may be space itself in the case of electromagnetic imaging. The interacting electric and magnetic fields which comprise electromagnetic radiation can exist even in a vacuum.

Theoretical descriptions of the wave propagation are constructed by considering the physics of the medium and the disturbance to it. In many situations, when the strength of the radiation is not too great and the medium is free of hysteresis effects, the "restoring force" towards equilibrium in the medium can be considered as proportional to the "displacement" from equilibrium. The propagation of waves is then a linear phenomenon (ϕ1.2.1) and the linear descriptions which follow are appropriate. This is true of acoustic waves of moderate amplitude [Morse and Ingard 1968 ϕ6], and of electromagnetic waves in all of the situations considered in this thesis [Born and Wolf 1975 ϕ1.1.2].

Brief definitions of some terms describing the properties of media are listed below. "Properties" means only those properties

affecting wave propagation.

LINEARITY (amplitude invariance). A linear medium is one in which the constitutive relations [Born and Wolf 1975 ϕ1.1.2] between the "displacement" and the "restoring force" are linear (see two paragraphs back). A nonlinear medium is one in which they are not. Waves in a linear medium obey the superposition principle.

HOMOGENEITY (position invariance). A homogeneous medium is one in which the properties are the same everywhere. An inhomogeneous medium is one in which the properties vary with position.

ISOTROPY (direction invariance). A medium is isotropic if the properties at each point are the same for all directions of propagation, otherwise it is anisotropic. Materials exhibiting long-range order in their structure, e.g. crystals, are frequently anisotropic. Random substances such as gases are usually isotropic. Isotropy also applies to sources: an isotropic source or isotrope is one which radiates equally in all directions.

DISPERSION (frequency-variance). A medium is non-dispersive (over a given frequency range) if the speed of propagation is the same for radiations of all temporal frequencies within the range, otherwise it is dispersive. Dispersion of light arises from the interaction between the electromagnetic field and the atoms and molecules that comprise matter [Born and Wolf 1975 ϕ2.3.4]. Gases, being tenuous, are not highly dispersive. Liquids and solids, which are far denser, can be much more dispersive. Dispersion in glass is the reason why prisms can produce optical spectra from white light.

1.5.1 LONGITUDINAL AND TRANSVERSE WAVES, POLARISATION

Wave motions can be classified into two types, according to whether the medium's displacements from equilibrium occur in a direction parallel to (longitudinal waves) or perpendicular to (transverse waves) the direction of propagation of the wave (see Figure 1.3). This definition has significance for mechanical wave motions in material media, where the displacements of particles can be

observed. In three-dimensional space there are three mutually orthogonal directions, one of which may be taken as parallel to the direction of propagation and the other two perpendicular to it. Thus, for a given direction of propagation, longitudinal displacements are constrained to a line whereas transverse displacements occur in a plane (Figure 1.3c). Transverse waves possess the property of polarisation [Jordan and Balmain 1968 ϕ12.06] which describes the direction of the displacement in the plane.

A longitudinal displacement is describable by one scalar function of time, A(t), the amplitude of the displacement. A transverse displacement requires two scalar functions, or equivalently a two-dimensional vector function, to describe it. A common choice is $A_x(t)$ and $A_y(t)$, the two rectangular components of the displacement.

A propagation medium has, in general, quite different

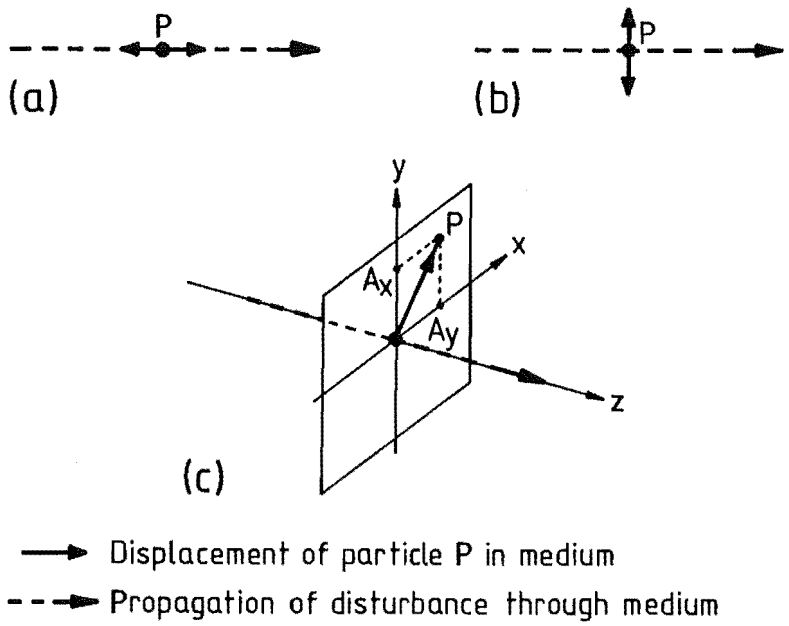


Figure 1.3: Transverse and longitudinal waves.

(a) Longitudinal wave: displacements caused by the wave are parallel to the direction of propagation of the wave.

(b) Transverse wave: displacements are perpendicular to the direction of propagation.

(c) Transverse wave in three dimensions: displacements can occur along any locus in a plane.

properties for longitudinal and transverse waves. For example, a fluid of negligible viscosity can support longitudinal (pressure) waves but not transverse waves because it cannot support shear stresses. Some media show different properties for different states of polarisation of transverse waves, e.g a calcite crystal, Polaroid film. Electromagnetic radiation was first shown to be a transverse wave when optical experiments demonstrated the polarisation of light. Strong [1958 ϕ6] briefly summarises the relevant history. Its transverse nature is also deducible theoretically from Maxwell's equations [Maxwell 1865, Born and Wolf 1975 ϕ1.4.1].

1.5.2 SCALAR DESCRIPTION OF WAVE MOTION

In this thesis, it is assumed that wave motions can be adequately described by a scalar quantity, representing the amplitude of a longitudinal wave or the amplitude of one component of a transverse wave. This requires the following conditions to be fulfilled (as they are in the physical situations examined herein):

1. The performance of the imaging system is independent of the polarisation of the radiation.

2. Differences in the behaviour of the electric and magnetic

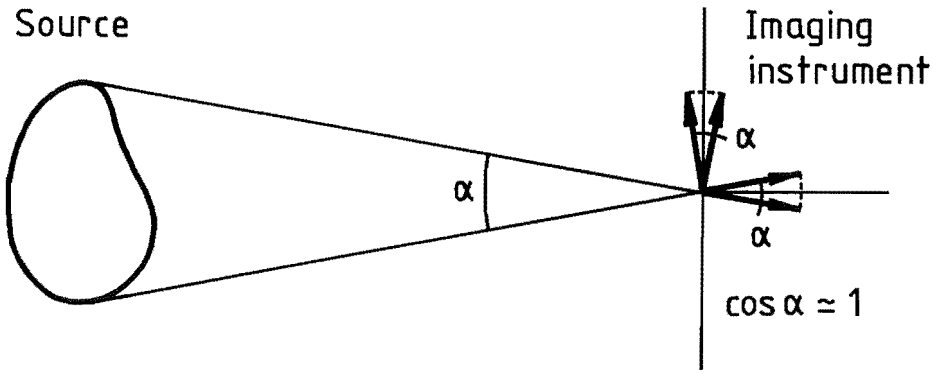


Figure 1.4: When the angle subtended by a source at the imaging instrument is small, the amplitude of the sum of the waves at the imaging instrument is approximately equal to the scalar sum of the individual wave amplitudes. The vector nature of the waves can then be ignored.

components of an electromagnetic field, as expressed through Maxwell's equations, can be ignored, so that considering one component of either the electric or the magnetic field yields sufficiently accurate results (see Born and Wolf [1975 ∅8.4]).

3. The field of view of the imaging system is small enough that rays emanating from different parts of the source and arriving at any particular point in the aperture plane of the imaging instrument can be taken as effectively parallel to each other, in the sense that $\cos(\alpha) \approx 1$ in Figure 1.4. Then the amplitude of the vector sum of the wave motions can be taken as being equal to the scalar sum of the amplitudes of the wave motions.

The scalar function describing the wave motions is called the wave function, $\Psi({}^3\underset{\sim}{x},t)$.

1.5.3 WAVE EQUATION FOR A HOMOGENEOUS MEDIUM

When the medium is linear, homogeneous, isotropic, and non-dispersive, the scalar wave function $\Psi({}^3\underset{\sim}{x},t)$ introduced in Section 1.5.2 satisfies the linear wave equation [cf. Jones 1964 ∅1.16]

$$\nabla^2\Psi({}^3\underset{\sim}{x},t) - \frac{1}{c^2}\frac{\partial^2\Psi({}^3\underset{\sim}{x},t)}{\partial t^2} = -{}^3\Sigma({}^3\underset{\sim}{x},t). \qquad (1.41)$$

Here, ∇^2 is the Laplacian operator [Kreyszig 1979 pp.395-396],

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2},$$

c is the velocity of propagation of the wave, and ${}^3\Sigma$ is the source density giving rise to Ψ.

If the form of $\Psi({}^3\underset{\sim}{x},t)$ is such that its temporal Fourier transform exists, or can be defined in the limit (physical realisability of Ψ is a sufficient condition for the existence of its Fourier transform; see Bracewell [1978 pp.8-13]), then it can be characterised as completely by its space-frequency dependence $\psi({}^3\underset{\sim}{x},\nu)$ as it is by its space-time dependence. Taking the temporal Fourier transform of (1.41), and using the derivative theorem for the Fourier

transform [Bracewell 1978], yields the space-frequency wave equation known as the Helmholtz equation,

$$\nabla^2\psi({}^3\underset{\sim}{x},\nu) + k^2\psi({}^3\underset{\sim}{x},\nu) = -{}^3\sigma({}^3\underset{\sim}{x},\nu), \tag{1.42}$$

where the wave number or propagation constant k is defined by

$$k = \frac{2\pi\nu}{c} = \frac{2\pi}{\lambda} \tag{1.43}$$

with ν being the temporal frequency and λ the wavelength, and ${}^3\sigma$ is the space-frequency representation of the source density ${}^3\Sigma$. The refractive index n of the medium, relative to some reference medium, is defined by

$$n = \frac{c_0}{c} = \frac{k}{k_0} \tag{1.44}$$

where c_0 is the speed of propagation in the reference medium and k_0 is the associated wave number. For electromagnetic radiation the reference medium is usually free space, i.e. a vacuum [Born and Wolf 1975 φ1.2]. The refractive index is then given by

$$n = \sqrt{\varepsilon_r \mu_r} ,$$

where ε_r and μ_r are the relative permittivity and relative permeability, respectively, of the medium. The great majority of materials (those which are not ferromagnetic) have a relative permeability very close to unity, so the refractive index can usually be taken as depending on the permittivity only, thus giving

$$n = \sqrt{\varepsilon_r} . \tag{1.45}$$

In dispersive materials, ε_r depends on the frequency ν.

The integral representation of the Helmholtz equation (1.42) is [cf. Jones 1964 φ1.16]

$$\psi({}^3\underset{\sim}{x},\nu) = \int_{{}^3T_S} {}^3\sigma({}^3\underset{\sim}{x}',\nu)\,\frac{e^{-jkr}}{4\pi r}\,.\,d^3\underset{\sim}{x}' \tag{1.46}$$

where

$$r = |{}^3\underset{\sim}{x} - {}^3\underset{\sim}{x}'|.$$

Equation (1.46) is a superposition integral (ϕ1.2.1) over the contributions to ψ from each part of the source ${}^3\sigma$. The term

$$\frac{e^{-jkr}}{4\pi r}$$

is called the free-space Green's function, and describes the expanding spherical wave from a point source of unit strength.

In source-free regions of the propagation medium, ${}^3\Sigma$ is zero and so the right hand sides of (1.41) and (1.42) become zero, i.e.

$$\nabla^2\psi({}^3\underset{\sim}{x},t) - \frac{1}{c^2}\frac{\partial^2\psi({}^3\underset{\sim}{x},t)}{\partial t^2} = 0, \tag{1.47}$$

$$\nabla^2\psi({}^3\underset{\sim}{x},\nu) + k^2\,\psi({}^3\underset{\sim}{x},\nu) = 0. \tag{1.48}$$

These equations are "homogeneous" in the mathematical sense [Kreyszig 1979 ϕ2.1], which is a different use of the term quite unconnected with the properties of the medium. Table 1.1 gives two important special solutions of (1.47) and (1.48). An important property of these solutions is that the shape of a wavefront does not change as it propagates: plane waves remain planar and spherical waves remain spherical in a homogeneous medium. The associated rays are straight.

Absorption in the medium can be described by making the wavenumber k complex, i.e.

$$k = k_r - jk_i \tag{1.49}$$

The monochromatic plane wave solution, for example, then becomes (from Table 1.1)

$$\begin{aligned}\psi({}^3\underset{\sim}{x},\nu) &= \psi_0\, e^{-j(k_r - jk_i)\,{}^3\underset{\sim}{x}\cdot{}^3\underset{\sim}{n}} \\ &= \psi_0\, e^{-jk_r\,{}^3\underset{\sim}{x}\cdot{}^3\underset{\sim}{n}}\, e^{-k_i\,{}^3\underset{\sim}{x}\cdot{}^3\underset{\sim}{n}}.\end{aligned} \tag{1.50}$$

The real exponential term in the second line of (1.50) represents an exponential loss in amplitude of the wave as it propagates in the $^3\underset{\sim}{n}$ direction, according to the absorption factor k_i.

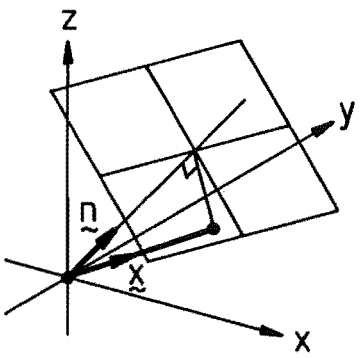
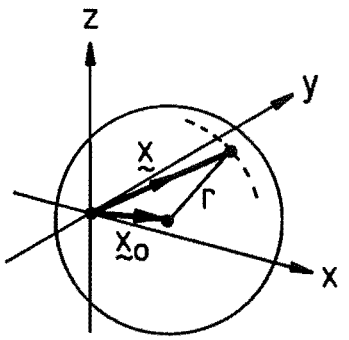
SOLUTION	PLANE WAVE	SPHERICAL WAVE $\underset{\sim}{r} = \underset{\sim}{x} - \underset{\sim}{x}_0, \quad r = \lvert\underset{\sim}{r}\rvert$
SPACE-TIME	$\Psi(\underset{\sim}{x},t) = \Psi_0\, e^{-jk\underset{\sim}{x}.\underset{\sim}{n}}\, e^{jkct}$	$\Psi(\underset{\sim}{x},t) = \Sigma_0 \dfrac{e^{-jkr}}{4\pi r} e^{jkct}$
SPACE-FREQUENCY	$\psi(\underset{\sim}{x},\nu) = \psi_0\, e^{-jk\underset{\sim}{x}.\underset{\sim}{n}}$	$\psi(\underset{\sim}{x},\nu) = \Sigma_0 \dfrac{e^{-jkr}}{4\pi r}$
DESCRIPTION	Monochromatic plane wave propagating in the direction of $\underset{\sim}{n}$ with speed c	Monochromatic spherical wave expanding from the point $\underset{\sim}{x}_0$ with speed c
GEOMETRY		
SOURCE	Infinite sheet, all points radiating in phase	Point source of strength Σ_0 at $\underset{\sim}{x} = \underset{\sim}{x}_0$
WAVEFRONTS	Parallel planes, normal to $\underset{\sim}{n}$	Concentric spheres, centred on $\underset{\sim}{x}_0$
RAYS	Parallel rays, parallel to $\underset{\sim}{n}$	Radially in all directions from $\underset{\sim}{x}_0$

Table 1.1: Two important special solutions in three-dimensional space of the wave equation for a source-free homogeneous medium.

1.5.3.1 Field at the Aperture Plane

Referring to the geometry of Figure 1.5, the Rayleigh-Sommerfeld formulation of scalar diffraction theory [cf. Goodman 1968 ϕ3.4] describes the wave motion at the aperture plane, ψ_1, in terms of that at the object plane, ψ_0, by the relation

$$\psi_1({}^2\underset{\sim}{\xi},\nu) = \frac{j}{\lambda}\int_{{}^2T_S} \psi_0({}^2\underset{\sim}{x},\nu)\,\frac{e^{-jkr}}{r}\cos({}^3\underset{\sim}{z},{}^3\underset{\sim}{r})\;.d\underset{\sim}{x} \qquad (1.51)$$

Here, ${}^2\underset{\sim}{x}$ is the two-dimensional position vector in the object plane, ${}^2\underset{\sim}{\xi}$ is the two-dimensional position vector in the aperture plane, and ${}^3\underset{\sim}{r}$ is the vector joining the ends of ${}^2\underset{\sim}{x}$ and ${}^2\underset{\sim}{\xi}$. r is the length of ${}^3\underset{\sim}{r}$. $\cos({}^3\underset{\sim}{z},{}^3\underset{\sim}{r})$ is the cosine of the angle between ${}^3\underset{\sim}{r}$ and the z-axis. z_0 is the distance between the object and aperture planes. ${}^2T_S$ is the region of the object plane outside which ψ_0 can be taken as zero, i.e. any radiation from the source ${}^3\sigma$ which passes through the object plane outside ${}^2T_S$ does not contribute to the image (see ϕ1.4.1).

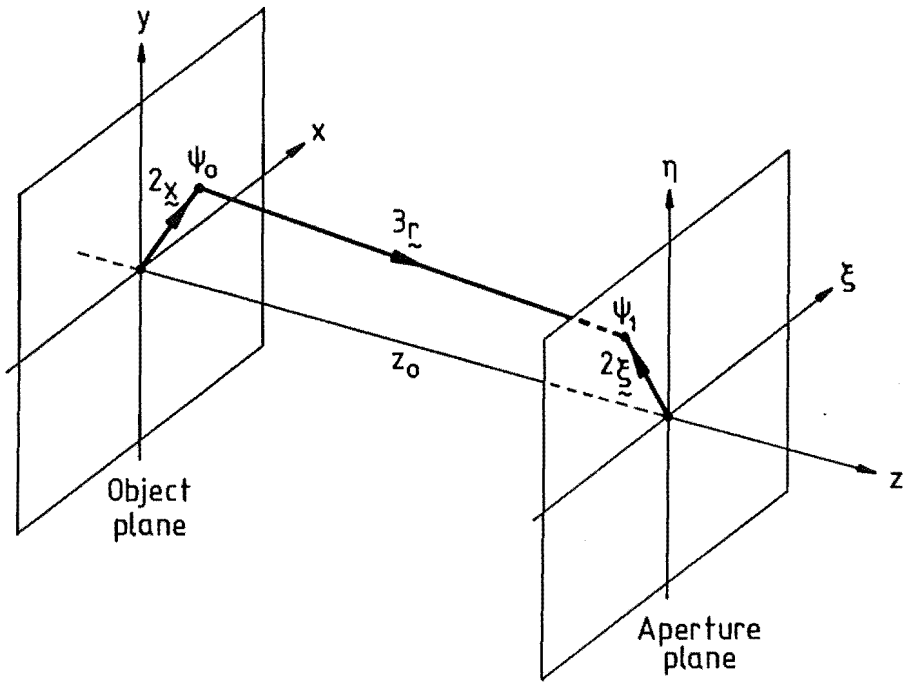


Figure 1.5: Geometry for calculating the field arising at the aperture plane from sources in the object plane (ϕϕ 1.5.3.1, 1.5.3.2).

Comparison of (1.51) with (1.46) suggests that (1.51) can be rewritten as

$$\psi_1({}^2\underset{\sim}{\xi},\nu) = \int_{{}^2T_S} {}^2\sigma({}^2\underset{\sim}{x},\nu)\,\frac{e^{-jkr}}{4\pi r}\,.d\underset{\sim}{x} \tag{1.52}$$

with ${}^2\sigma$ being the equivalent source of Section 1.4.1. Each temporal frequency component of the equivalent source ${}^2\sigma$ is thus proportional to the same component of the object plane field ψ_0 with

(a) a frequency-dependent scaling factor $\frac{1}{\lambda}$,

(b) a phase lead of 90°, and

(c) the obliquity factor $\cos({}^3\underset{\sim}{x},{}^3\underset{\sim}{r})$, i.e. each part of the equivalent source radiates nonisotropically.

1.5.3.2 Fresnel and Fraunhofer Diffraction

When the distance z_0 between the object plane and the aperture plane is much larger than the maximum dimension of the equivalent source region ${}^2T_S$ and the diameter D of the imaging instrument's aperture, (1.51) can be simplified usefully to good approximation. Two particular approximations, the Fresnel approximation and the Fraunhofer approximation, are discussed below.

The obliquity factor in (1.51) is

$$\cos({}^3\underset{\sim}{z},{}^3\underset{\sim}{r}) = \cos\left(\frac{1}{\sqrt{1+(|{}^2\underset{\sim}{\xi}-{}^2\underset{\sim}{x}|^2/z_0)^2}}\right). \tag{1.53}$$

This can be approximated by

$$\cos({}^3\underset{\sim}{z},{}^3\underset{\sim}{r}) \approx 1 \tag{1.54}$$

when the angle $({}^3\underset{\sim}{z},{}^3\underset{\sim}{r})$ is small, incurring for example an error of less than 6% when the angle is less than 20°. Similarly, when $({}^3\underset{\sim}{z},{}^3\underset{\sim}{r})$ is small, one has $r \approx z_0$ and so may write

$$\frac{e^{-jkr}}{r} \approx \frac{e^{-jkr}}{z_0}\,. \tag{1.55}$$

The approximation $r = z_0$ cannot be used in the exponent in (1.55) because it is multiplied by the potentially very large number k, leading to phase errors of many times 2π unless the angle $({}^3\underset{\sim}{z}, {}^3\underset{\sim}{r})$ is exceedingly small.

The Fresnel approximation and the Fraunhofer approximation are each approximations to r in the exponent of (1.55). The exact expression for r is

$$r = \sqrt{z_0 + |{}^2\underset{\sim}{\xi} - {}^2\underset{\sim}{x}|^2}$$
$$= z_0 \sqrt{1 + (|{}^2\underset{\sim}{\xi} - {}^2\underset{\sim}{x}|^2 / z_0)^2}\ , \qquad (1.56)$$

which can be expanded using the binomial series, if $|{}^2\underset{\sim}{\xi} - {}^2\underset{\sim}{x}| < z_0$, to give

$$r = z_0 \left[1 + \frac{1}{2}(|{}^2\underset{\sim}{\xi} - {}^2\underset{\sim}{x}|/z_0)^2 - \frac{1}{8}(|{}^2\underset{\sim}{\xi} - {}^2\underset{\sim}{x}|/z_0)^4 + \ldots \right]. \qquad (1.57)$$

The Fresnel approximation retains only the first two terms of the binomial series, yielding

$$r \approx z_0 \left[1 + \frac{|{}^2\underset{\sim}{\xi} - {}^2\underset{\sim}{x}|^2}{2z_0^2} \right] \qquad (1.58)$$

in (1.55) and hence, from (1.51) and (1.54),

$$\psi_1({}^2\underset{\sim}{\xi}, \nu) \approx \frac{j}{\lambda} \int_T \psi_0({}^2\underset{\sim}{x}, \nu)\, \frac{e^{-jkz_0}}{z_0}\, e^{-jk\frac{1}{2z_0}|{}^2\underset{\sim}{\xi} - {}^2\underset{\sim}{x}|^2} .d^2\underset{\sim}{x}$$
$$= \frac{j}{\lambda z_0} e^{-jkz_0}\, e^{-jk\frac{|{}^2\underset{\sim}{\xi}|^2}{2z_0}} \int_T \psi_0({}^2\underset{\sim}{x}, \nu)\, e^{-jk\frac{|{}^2\underset{\sim}{x}|^2}{2z_0}}\, e^{j\frac{2\pi}{\lambda z_0}\, {}^2\underset{\sim}{\xi}.\,{}^2\underset{\sim}{x}} .d^2\underset{\sim}{x}\ . \qquad (1.59)$$

This expression, apart from amplitude and phase factors outside the integral on the second line which are independent of ${}^2\underset{\sim}{x}$, is a Fourier transform of the object-plane field distribution over ${}^2T_S$ multiplied by a quadratic phase factor $e^{-jk|{}^2\underset{\sim}{x}|^2/(2z_0)}$, with the Fourier transform variable ${}^2\underset{\sim}{u} = \frac{{}^2\underset{\sim}{\xi}}{\lambda z_0}$ (cf. 1.12b).

The Fraunhofer approximation imposes a more stringent condition

$$2z_0 >> k \left|{}^2\underset{\sim}{x}\right|^2_{max} \tag{1.60}$$

on the size of the source region relative to z_0. This allows the quadratic phase factor within the integral on the second line of (1.59) to be approximated by unity over the source region, yielding the result

$$\psi_1({}^2\underset{\sim}{\xi},\nu) = \frac{j}{\lambda z_0} e^{-jkz_0} e^{-jk\frac{\left|{}^2\underset{\sim}{\xi}\right|^2}{2z_0}} \int_{{}^2T_S} \psi_0({}^2\underset{\sim}{x},\nu)\, e^{j\frac{2\pi\, {}^2\underset{\sim}{\xi}.{}^2\underset{\sim}{x}}{\lambda z_0}} .d^2\underset{\sim}{x} . \tag{1.61}$$

Apart from the multiplicative amplitude and phase factors outside the integral, the Fraunhofer diffraction pattern $\psi_1({}^2\underset{\sim}{x},\nu)$ is the spatial Fourier transform of the object plane field distribution over ${}^2T_S$, and hence of the equivalent source (φ1.4.1), again with the scaled frequency coordinates ${}^2\underset{\sim}{u} = \frac{{}^2\underset{\sim}{\xi}}{\lambda z_0}$. The Fourier transform integral in (1.61) is sometimes referred to as the angular spectrum of plane waves, because (cf. Table 1.1 and (1.43)) the integral represents a superposition of plane waves propagating in the directions ${}^2\underset{\sim}{\alpha} = \frac{{}^2\underset{\sim}{x}}{z_0}$.

Since (1.61) is a simplification of (1.59), (1.59) is applicable wherever (1.61) is and so Fraunhofer diffraction is a special case of Fresnel diffraction. The range of z_0 over which (1.61) is applicable defines the Fraunhofer region or far field of the source. The range of z_0 over which (1.59) is applicable but (1.61) is not defines the Fresnel region or near field of the source. The transition from the near field to the far field is gradual, but a common rule of thumb defines the far field by [Silver 1965 φ6.9]

$$z_0 > \frac{2E^2}{\lambda} \tag{1.62}$$

where E is the maximum linear dimension of ${}^2T_S$. Distances to the far field according to this formula are tabulated for various values of E in Table 1.2, for a wavelength equal to that of green light in free space.

Maximum linear dimension E	Distance to far field $z_0 = \frac{2E^2}{\lambda}$ (green light, λ = 550nm)
20 μm	1.5 mm
100 μm	36 mm
1 mm	3.6 m
1 cm	360 m
10 cm	36 km
1 m	3600 km
5 m	91000 km

Table 1.2: Distance from a source or aperture to its far field, according to the rule of thumb (1.62).

The Fraunhofer diffraction pattern of a source can be observed over a limited region much closer to the source than given by (1.62) (i.e. the far field can be brought near) if the source lies not on a plane but on a sphere (e.g. the celestial sphere of astronomy) centred on the observing region, or (in coherent optical computing where the source is often a transparency) if the source is illuminated by spherical waves converging towards the observing region.

1.5.4 WAVE EQUATION FOR AN INHOMOGENEOUS MEDIUM

When the refractive index of the medium varies as a function of position, the propagation of waves is different from that in a homogeneous medium. Plane waves propagating through the medium do not remain planar, and rays instead of being straight become bent. If the variations in refractive index are small over distances equal to the longest wavelength of the radiation of interest, wave propagation in source-free regions of the inhomogeneous medium can be described by the wave equation [Ishimaru 1978 ϕ17.1]

$$\nabla^2 \psi({}^3\underset{\sim}{x},\nu) + k^2({}^3\underset{\sim}{x})\, \psi({}^3\underset{\sim}{x},\nu) = 0 \qquad (1.63)$$

of which (1.48) is a special case.

Following Ishimaru (1978), the refractive index can be written as the sum of its average value $\bar{n}$ and a relative fluctuation n_1,

$$n(^3\underset{\sim}{x}) = \bar{n}\ (1+n_1(^3\underset{\sim}{x})). \tag{1.64}$$

Defining $\bar{k} = k_0\bar{n}$ (cf. 1.44) gives $k(^3\underset{\sim}{x}) = \bar{k}\ (1+n_1(^3\underset{\sim}{x}))$, which substituted into (1.63) yields

$$\nabla^2\psi(^3\underset{\sim}{x},\nu) + \bar{k}^2\ (1+n_1(^3\underset{\sim}{x}))^2\ \psi(^3\underset{\sim}{x},\nu) = 0. \tag{1.65}$$

This may be rearranged to put the fluctuations on the right hand side, giving

$$\nabla^2\psi(^3\underset{\sim}{x},\nu) + \bar{k}^2\ \psi(^3\underset{\sim}{x},\nu) = -\bar{k}^2\ \delta n(^3\underset{\sim}{x})\ \psi(^3\underset{\sim}{x},\nu) \tag{1.66}$$

where $\delta n = (1+n_1)^2 - 1 = 2n_1 + n_1^2$. Note that there are no terms on the left hand side of (1.66) describing the inhomogeneity of the medium. (1.66) has the same form as the wave equation for a homogeneous medium (1.42), with the inhomogeneities appearing in the guise of sources. These fictitious sources are known as <u>induced</u> or <u>polarisation</u> sources [Bates and Ng 1972]. They differ from true sources in that they depend on the field ψ.

The total field ψ can be regarded as the sum of the field ψ_o which would exist in the absence of the fluctuations and the field ψ_s arising from the induced sources, i.e.

$$\psi = \psi_o + \psi_s\ . \tag{1.67}$$

ψ_o and ψ_s may be called the <u>incident</u> and the <u>scattered</u> fields respectively, the terminology coming from the remote probing problem [Bates et al 1983] in which an incident field is externally applied to a region of inhomogeneity, and information about the inhomogeneity is obtained from measurements of the field scattered by it (see Figure 1.6). In this situation the inhomogeneities in the medium are termed <u>scatterers</u> (cf. first paragraph of ϕ1.4) and occur within the <u>scattering region</u> $^3\psi_{sc}$.

The integral representation of (1.66) is (cf. (1.46), in which there is no externally applied "incident" field)

$$\psi({}^3\underset{\sim}{x},\nu) = \psi_o({}^3\underset{\sim}{x},\nu) + \bar{k}^2 \int_{{}^3T_{sc}} \delta n({}^3\underset{\sim}{x}')\, \psi({}^3\underset{\sim}{x}',\nu)\, \frac{e^{-j\bar{k}r}}{4\pi r}\, .d^3\underset{\sim}{x}' \ , \tag{1.68}$$

with $r = |{}^3\underset{\sim}{x} - {}^3\underset{\sim}{x}'|$ as before. Exact solution of (1.68) is in general not practical, so approximate solutions are usually sought. Three approximate solutions based on perturbation methods [Morse and Feshbach 1953 ϕ9] are described below.

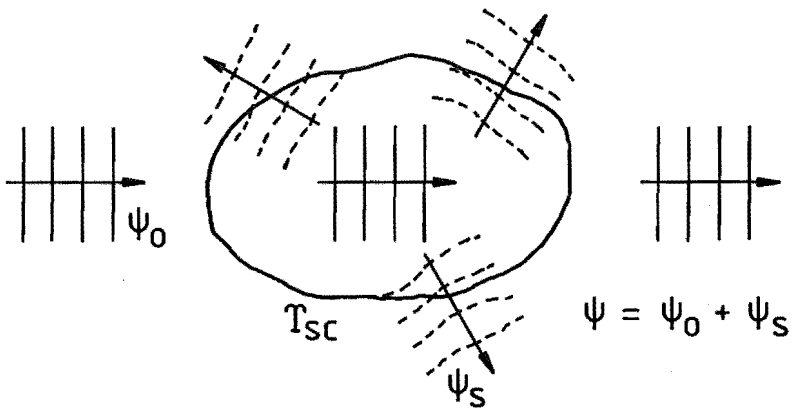


Figure 1.6: Inhomogeneous medium in the remote probing situation. An unperturbed field ψ_o incident on inhomogeneities within a scattering region T_{sc} gives rise to a scattered field ψ_s. The sum of ψ_o and ψ_s is the perturbed total field ψ.

1.5.4.1 The Born Approximation

The method of successive approximations can be used to solve (1.68) by writing the total field as a series

$$\psi({}^3\underset{\sim}{x},\nu) = \sum_{n=0}^{\infty} \psi_n({}^3\underset{\sim}{x},\nu) \tag{1.69}$$

where ψ_0 is usually, but not necessarily [Beylkin and Oristaglio

1985], taken to be the unperturbed incident field ψ_o, and

$$\psi_n(^3\underset{\sim}{x},\nu) = \bar{k}^2 \int_{^3T_{sc}} \delta n(^3\underset{\sim}{x}') \, \psi_{n-1}(^3\underset{\sim}{x}',\nu) \, \frac{e^{-j\bar{k}r}}{4\pi r} \, .d^3\underset{\sim}{x}' \, . \qquad (1.70)$$

This series converges only if the value of the largest eigenvalue associated with (1.70) is less than unity [Kreyszig 1979 φ19.10]. The terms ψ_n for n = 1,2,... represent the contributions to the scattered field of first-order scattering, second-order scattering, etc.

Born's first approximation, or simply the Born approximation, to the total field is obtained by taking only the first two terms in (1.69). This is equivalent to inserting ψ_o in place of ψ in the right hand side of (1.68) [Morse and Feshbach 1953 p.1073]. By ignoring all terms of order higher than 1, the Born approximation accounts for only first order scattering (interaction of the incident field with the inhomogeneities) and ignores multiple scattering (interaction of the scattered field with the inhomogeneities). It is accurate provided that both the magnitude and the phase of the total field ψ are similar to those of the incident field ψ_o at all points in the scattering region. The magnitude similarity requires $|\psi_s| << |\psi_o|$, which implies that

$$|\delta n| << 1, \qquad (1.71)$$

i.e. the fluctuations are weak. The phase similarity requires that the change in effective path length (i.e. integrated refractive index fluctuation) to each point in $^3T_{sc}$, due to the inhomogeneity, is small. This places a combined constraint on the severity of the fluctuations and the size of $^3T_{sc}$, which is expressible by the relation

$$\bar{k} \, L \, |n_1|_{max} << 2\pi \qquad (1.72)$$

where L is the largest linear dimension of $^3T_{sc}$.

1.5.4.2 The Rytov Approximation

The Rytov approximation [Chernov 1967 ϕ16] is obtained by writing the field as

$$\psi(^3\underline{x},\nu) = e^{\Upsilon(^3\underline{x},\nu)} , \tag{1.73}$$

and then attempting to solve for the transformed quantity Υ. Substituting (1.73) into (1.66) yields the Riccati equation [Ishimaru 1978 ϕ17.2.2]

$$\nabla^2\Upsilon + \nabla\Upsilon.\nabla\Upsilon + \bar{k}^2 (1+\delta n) = 0. \tag{1.74}$$

If Υ is written as the sum of an "incident" and a "scattered" part,

$$\Upsilon = \Upsilon_o + \Upsilon_s , \tag{1.75}$$

where Υ_o can be taken as being related to the incident field ψ_o by

$$\psi_o = e^{\Upsilon_o}, \tag{1.76}$$

the Riccati equation can be manipulated [Ishimaru 1978] to give

$$\nabla^2(\psi_o\Upsilon_s) + \bar{k}^2 \psi_o \Upsilon_s = -(\nabla\Upsilon_s.\nabla\Upsilon_s + \bar{k}^2\delta n) \psi_o. \tag{1.77}$$

This is an inhomogeneous wave equation for $\psi_o\Upsilon_s$, whose integral representation (cf. 1.46) yields

$$\Upsilon_s(^3\underline{x},\nu) = \frac{1}{\psi_o} \int_{^3T_{sc}} (\nabla\Upsilon_s.\nabla\Upsilon_s + \bar{k}^2\delta n) \psi_o \frac{e^{-j\bar{k}r}}{4\pi r} .d^3\underline{x}' . \tag{1.78}$$

The Rytov approximation to Υ is obtained by setting $\nabla\Upsilon_s = 0$ in (1.78), giving the first-order approximation to Υ_s,

$$\Upsilon_1(^3\underline{x},\nu) = \frac{1}{\psi_o} \int_{^3T_{sc}} \bar{k}^2 \delta n \psi_o \frac{e^{-j\bar{k}r}}{4\pi r} .d^3\underline{x}' . \tag{1.79}$$

Higher order approximations can be obtained by writing Υ as a series

$$\Upsilon({}^3\underset{\sim}{x},\nu) = \sum_{n=0}^{\infty} \Upsilon_n({}^3\underset{\sim}{x},\nu) \tag{1.80}$$

with Υ_0 equal to Υ_o (1.76), Υ_1 given by (1.79), and the higher order terms given by (from (1.78) and 1.79))

$$\Upsilon_n({}^3\underset{\sim}{x},\nu) = \frac{1}{\psi_o} \int_{{}^3T_{sc}} (\nabla\Upsilon_{n-1}.\nabla\Upsilon_{n-1})\ \psi_o\ \frac{e^{-j\bar{k}r}}{4\pi r}\ .d^3\underset{\sim}{x}' \tag{1.81}$$

for $n > 1$ [cf. Ishimaru 1978].

The Rytov approximation is written in terms of the field ψ as

$$\begin{aligned}\psi_r({}^3\underset{\sim}{x},\nu) &= e^{\Upsilon_o + \Upsilon_1} \\ &= \psi_o({}^3\underset{\sim}{x},\nu)\ e^{\Upsilon_1({}^3\underset{\sim}{x},\nu)}\ .\end{aligned} \tag{1.82}$$

If the exponential in (1.82) is expanded as a power series, taking the first two terms gives $\psi_1 = \psi_o(1 + \Upsilon_1)$ which is identical to the Born approximation.

The Rytov approximation (1.79) is valid if

$$|\nabla\Upsilon_s\ /\ k| \ll 1 \tag{1.83}$$

which means that the amount of scattering over the distance of one wavelength is small [Chernov 1967]. No restriction analogous to (1.72) is imposed on the size of the scattering region.

The Rytov approximation is generally regarded as superior to the Born approximation, especially when the inhomogeneous region is large but the inhomogeneities themselves are "smooth" on the scale of a wavelength (e.g. for optical propagation through the Earth's atmosphere) so that scattering is predominantly in the forward direction [Ishimaru 1978 ϕ6.6]. The Born approximation is more applicable to widely separated inhomogeneities of small size which

scatter significantly in all directions.

The real part of γ in (1.73) is known as the log-modulus of the field, and its imaginary part is the phase of the field. The statistics of these quantities are significant when considering propagation through random, turbulent inhomogeneities (ϕ2.3.1).

1.5.4.3 An Extended Rytov Approximation

Bates et al [1976b] have proposed an extension to the Rytov approximation that partly allows for the gradient terms dropped in going from (1.78) from (1.79). When the inhomogeneities are such that refraction effects predominate over reflection, the radiation is well described by the approximate relation

$$\psi({}^3\underset{\sim}{x},\nu) \approx n^{-\frac{1}{2}} e^{-jk\int_{\tau_0}^{\tau} n.ds} \tag{1.84}$$

which represents a tube of rays that propagate along paths curved by the inhomogeneities [Felsen and Marcuvitz 1973 ϕ1.7]. The $n^{-\frac{1}{2}}$ term describes changes in amplitude as the velocity changes, and the factor multiplying -j in the exponent represents the phase delay along each ray τ from a starting point τ_0, with s being path length along the ray. The corresponding formulation for the unperturbed case (constant n) is

$$\psi_o({}^3\underset{\sim}{x},\nu) = n^{-\frac{1}{2}} e^{-jk(\tau-\tau_0)n}$$

which describes a tube of straight rays of constant amplitude.

From (1.73), (1.75), and (1.84), an approximate expression

$$\begin{aligned}\nabla\gamma_s &\approx -jk(\underset{\sim}{\tau}'n-\underset{\sim}{\tau}) - \tfrac{1}{2}\nabla\ln(n)\\ &\approx -jk\underset{\sim}{\tau}(n-1) - \tfrac{1}{2}\nabla\ln(n)\end{aligned} \tag{1.85}$$

is obtained, where $\underset{\sim}{\tau}'$ and $\underset{\sim}{\tau}$ are unit vectors describing the forward directions of the curved rays (ψ) and the straight rays (ψ_o) respectively, and the ray curvature is assumed to be small so that

$\underline{\tau}'.\underline{\tau} \approx 1$. The estimate of $\nabla\gamma_s$ thus obtained can be inserted into (1.78) to obtain an improved approximation to γ_s.

Numerical results [Dunlop et al 1976] confirm that the extended Rytov approximation is more accurate than the standard Born or Rytov approximations for forward scattering.

1.6 THE IMAGING INSTRUMENT

The imaging instrument is designed to produce an image from the radiation incident upon its aperture. The archetypal imaging instrument is the thin lens [Goodman 1968 ϕ5]. Practical imaging instruments usually have several or many components, e.g. multiple lenses or mirrors in an optical telescope, or they may produce an image from aperture plane measurements synthetically within a computer, but their ideal image-forming properties can be discussed in terms of the thin lens.

1.6.1 IDEAL IMAGING PROPERTIES OF THE THIN LENS

A (converging) lens is a piece of transparent material having a refractive index different from that of the surrounding medium, shaped so that the delay experienced by waves passing through the lens is greatest at its centre and reduces towards its edges in such a manner that diverging spherical wavefronts are transformed into converging spherical wavefronts by their passage through the lens (see Figure 1.7a). A useful abstraction is to consider the delays as occurring entirely in the plane through the centre of the lens, so that radiation emerges from the lens with the same transverse position as where it enters and experiences a delay independent of its direction of propagation through the lens. A lens which behaves in approximately this way, to a sufficient accuracy, is known as a thin lens. The ideal imaging properties discussed in this section are those of an appropriately figured thin lens.

Waves diverging from the point P_0 in Figure 1.7a, after travelling through the lens, converge (provided P_0 is not too close to

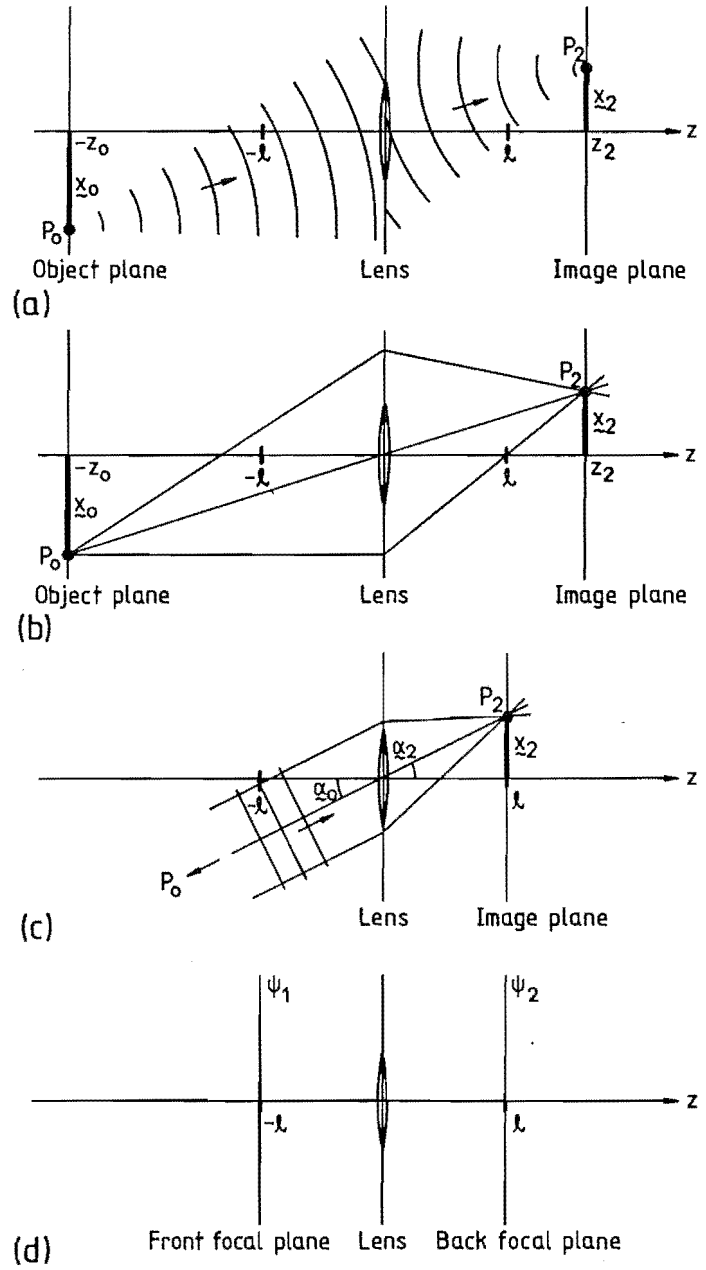


Figure 1.7: Imaging properties of the ideal thin lens.

(a) Wavefronts: spherical wavefronts diverging from P_0 become spherical wavefronts converging to P_2.

(b) Rays: all rays passing through P_0 and the lens pass through P_2.

(c) Object at infinity: angular position in the object is imaged as linear position in the image plane.

(d) Fourier transform relationships: the field ψ_2 in the back focal plane is the spatial Fourier transform of the field ψ_1 in the front focal plane.

the lens) to a point P_2. Thus an image of the point P_0 is formed at P_2 [Newton 1704 Book 1 Axiom VII]. The longitudinal distances z_0 and z_2 are related by the lens law

$$\frac{1}{z_0} + \frac{1}{z_2} = \frac{1}{\ell} \tag{1.86}$$

where ℓ is the focal length of the lens (ℓ is used in this thesis instead of the usual f to avoid confusion with the object f). The transverse distances $\underset{\sim}{x}_0$ and $\underset{\sim}{x}_2$ are related by the magnification

$$M = \frac{x_2}{x_0} = \frac{-z_2}{z_0} \tag{1.87}$$

where x_0 and x_2 are corresponding components of $\underset{\sim}{x}_0$ and $\underset{\sim}{x}_2$. In this way, an image of the object plane $z = z_0$ is formed in the image plane $z = z_2$. The position vectors $\underset{\sim}{x}_0$ and $\underset{\sim}{x}_2$ are two-dimensional in most optical imaging systems. However the superscripts used to denote dimensionality in Sections 1.4 and 1.5 have been dropped because the results of this section can be generalised to dimensionalities other than two, such as one-dimensional imaging (e.g. with cylindrical lenses).

The ray picture equivalent to Figure 1.7a is shown in Figure 1.7b. Rays passing through the lens are bent towards its centre, so that all rays from P_0 passing through the lens intersect at P_2. Rays passing through the centre of the lens are not bent.

When the object plane is removed to infinite distance, i.e. $z_0 \to \infty$, the spherical wavefronts from any particular point in the object are effectively planar when they reach the lens. The associated rays are parallel and arrive at an angle $\underset{\sim}{\alpha}_0$ to the lens's principal axis, where $\underset{\sim}{\alpha}_0$ is the angular position of the point in the object as seen from the lens. In this situation (see Figure 1.7c), the image is formed in the back focal plane of the lens at $z_2 = \ell$, which is consistent with setting $z_0 = \infty$ in (1.86). The transverse angle $\underset{\sim}{\alpha}_2$ in the image is equal to the object angular position $\underset{\sim}{\alpha}_0$, which can be seen by considering the unbent central ray in Figure 1.7c, so the angle-to-position magnification of the system is given by

$$M = \frac{x_2}{\alpha_0} \approx -\ell \tag{1.88}$$

when the angle $\underset{\sim}{\alpha}_0$ is small.

The mapping of angles to positions demonstrates the Fourier transforming property of the lens. The superposition of plane waves at the aperture plane, arriving at different angles from different parts of the object, is equivalent to the Fourier spectrum of the object (cf. the angular spectrum of plane waves, ϕ1.5.3.2). In forming an image, the lens is inverting the Fourier transform.

The Fourier transform relation between the field at the aperture plane of a lens and the field at the back focal plane is not exact, because the focal plane field exhibits an additional phase curvature [Goodman 1968 ϕ5.2]. This phase curvature is unimportant if the field is measured with a phase-insensitive detector such as an optical detector, or if it is measured over a spherical rather than a planar surface. An exact Fourier transform relationship between planes does, however, exist between the fields at the front and the back focal planes of the lens (see Figure 1.7d), in which special case the phase curvature disappears. This configuration is important in coherent optical computing, because it allows the two-dimensional Fourier transform to be performed easily and essentially instantaneously.

1.6.2 DIFFRACTION LIMITING

If the lens were of infinite size, then all radiation from the object plane propagating in the positive z-direction would be intercepted by the lens, and the image of a point source at $\underset{\sim}{x}_0$ in the object would be a point at $\underset{\sim}{x}_2 = M\,\underset{\sim}{x}_0$ in the image plane. If an equivalent source density $\sigma_0(\underset{\sim}{x}_0,\nu)$ were considered to lie in the object plane, then the field in the image plane would be given by

$$\psi_g(\underset{\sim}{x}_2,\nu) = \frac{1}{|M|}\,\sigma_0\left(\frac{-\underset{\sim}{x}_2}{M},\nu\right), \tag{1.89}$$

according to the predictions of geometrical optics [Goodman 1968 p.95]. In fact this is an unrealisable abstraction because a lens of infinite size cannot be constructed in reality.

Any real-world imaging system possesses an aperture of finite size, which is the region in the aperture plane outside which radiation incident upon the aperture plane does not contribute to the image. For a single thin lens, the lens plane is the aperture plane and the aperture is simply the size of the lens. Since the image plane field is formed by the interference of radiation from the whole aperture, diffraction effects associated with the finite aperture size cause the image of a point object to be spread out. The imagery obtained from an imaging instrument which is "perfect" except for a finite aperture is termed diffraction-limited. It represents the best performance that can reasonably be sought in practice (in the sense that it is more "reasonable" to imagine building an instrument with perfect imaging properties within its finite aperture than it is to imagine building one with an aperture of infinite size). The aperture is described by a pupil function

$$P(\underset{\sim}{\xi}) = \begin{cases} 1 & \text{for } \underset{\sim}{\xi} \text{ within the aperture} \\ 0 & \text{for } \underset{\sim}{\xi} \text{ outside it.} \end{cases} \tag{1.90}$$

The image plane field produced by a diffraction-limited imaging instrument is given by the convolution [cf. Goodman 1968 eq.(6.4)]

$$\psi_2(\underset{\sim}{x}_2,\nu) = \psi_g(\underset{\sim}{x}_2,\nu) \circledcirc h'(\underset{\sim}{x}_2,\nu) \tag{1.91}$$

where ψ_g is the image predicted by geometrical optics, given by (1.89), and the point spread function h′ is the Fraunhofer diffraction pattern from the aperture, given by

$$h'(\underset{\sim}{x}_2,\nu) = C \int_{-\infty}^{\infty} P(\underset{\sim}{\xi})\, e^{j\frac{2\pi}{\lambda z_2}\underset{\sim}{\xi}.\underset{\sim}{z}_2} \tag{1.92}$$

where C is a complex constant and $\underset{\sim}{\xi}$ is the spatial coordinate within the aperture plane.

1.6.3 COHERENT IMAGING

For coherent imaging, the image s_c is defined as the complex

amplitude of the field in the image plane, which is given by (1.91), and the object f_c is related to the complex amplitude of the field in the object plane by (1.94) below. An equation of the form of (1.11),

$$s_c(\underset{\sim}{x},\nu) = f_c(\underset{\sim}{x},\nu) \circledcirc h_c(\underset{\sim}{x},\nu) \tag{1.93}$$

can be obtained from (1.91) and (1.92) by the transformations

$$\begin{aligned} s_c(\underset{\sim}{x},\nu) &= \psi_2(\underset{\sim}{x},\nu), \\ f_c(\underset{\sim}{x},\nu) &= C\ (\lambda z_2)^K\ \psi_g(\underset{\sim}{x},\nu), \\ h_c(\underset{\sim}{x},\nu) &= h'(\underset{\sim}{x},\nu)\ /\ C\ (\lambda z_2)^K \end{aligned} \tag{1.94}$$

where C is the complex constant in (1.92) and K is the dimensionality of the position vectors $\underset{\sim}{x}$, $\underset{\sim}{x}_2$, $\underset{\sim}{\xi}$ in (1.92) and (1.94). When the object plane is at infinite distance from the aperture plane, $z_2 = \ell$ as noted previously (Figure 1.7c). Taking the spatial Fourier transform (1.12a) of (1.93) yields the Fourier-domain relation

$$S_c(\underset{\sim}{u},\nu) = F_c(\underset{\sim}{u},\nu)\ H_c(\underset{\sim}{u},\nu)\ . \tag{1.95}$$

From (1.94) and (1.92) the transfer function is seen to be

$$H_c(\underset{\sim}{u},\nu) = P(\lambda z_2 \underset{\sim}{u})\ , \tag{1.96a}$$

with

$$u = \frac{\xi}{\lambda z_2}\ . \tag{1.96b}$$

$\underset{\sim}{x}$ has units of distance in the image plane, and the spatial frequency $\underset{\sim}{u}$ has units of cycles per unit length in the image plane. The quantity $F_c(\underset{\sim}{u},\nu)$ may be (loosely) considered as the field at the aperture plane of the imaging instrument (see the final paragraph of ϕ1.6.1).

In optical imaging systems, available detectors respond to the intensity of the field, which is a time average of power over many cycles of the radiation. For these systems, "coherent" imaging requires that the object be spatially coherent (ϕ1.4.2.2) so that the

system appears linear in complex amplitude over the averaging time of the detector [Goodman 1968 ϕ6.1].

1.6.4 INCOHERENT IMAGING

For incoherent imaging, the source is spatially incoherent (ϕ1.4.2.2) and the image s_i is defined as the intensity of the field in the image plane, expressed as

$$s_i(\underset{\sim}{u},\nu) = \langle|\psi_2(\underset{\sim}{x},\nu)|^2\rangle = \langle|s_c(\underset{\sim}{x},\nu)|^2\rangle \tag{1.97}$$

where the angle brackets denote a time average over an interval subject to the same constraints as given in Section 1.4.2. The intensity is, for example, the property of the optical field measured by all known optical detectors. Substituting (1.93) into (1.97) gives

$$\begin{aligned} s_i(\underset{\sim}{x},\nu) &= \langle|f_c(\underset{\sim}{x},\nu) \circledcirc h_c(\underset{\sim}{x},\nu)|^2\rangle \\ &= \iint_{-\infty}^{\infty} \langle f_c^*(\underset{\sim}{x}',\nu) f_c(\underset{\sim}{x}'',\nu)\rangle\, h_c^*(\underset{\sim}{x}-\underset{\sim}{x}',\nu) h_c(\underset{\sim}{x}-\underset{\sim}{x}'',\nu)\, .d\underset{\sim}{x}'\, .d\underset{\sim}{x}'' \end{aligned} \tag{1.98}$$

after expanding the convolution and interchanging the order of the integrations and averaging. The quantity

$$\langle f_c^*(\underset{\sim}{x}',\nu) f_c(\underset{\sim}{x}'',\nu)\rangle$$

is the mutual intensity or mutual coherence of f_c (cf. 1.29a).

If the source giving rise to f_c is spatially incoherent, then f_c is itself spatially incoherent and its mutual intensity function may be described by (cf. 1.40)

$$\langle f_c^*(\underset{\sim}{x}',\nu) f_c(\underset{\sim}{x}'',\nu)\rangle = \langle f_c^*(\underset{\sim}{x}',\nu) f_c(\underset{\sim}{x}',\nu)\rangle\, \delta(\underset{\sim}{x}'-\underset{\sim}{x}'').$$

The notation $f_c(\underset{\sim}{x}',\nu)$ can now refer only to a quasi-monochromatic object (ϕ1.4.2.1) of mean temporal frequency ν, because a strictly monochromatic object cannot be spatially incoherent [cf. Born and Wolf

1975 ϕ10.4.2 and especially the footnote on p.513]. Substituting the latest equation above into (1.98) yields the result

$$s_i(\underset{\sim}{x},\nu) = \int_{-\infty}^{\infty} \langle f_c^*(\underset{\sim}{x}',\nu) f_c(\underset{\sim}{x}',\nu)\rangle \, h_c^*(\underset{\sim}{x}-\underset{\sim}{x}',\nu) h_c(\underset{\sim}{x}-\underset{\sim}{x}',\nu) \, .dx'$$

$$= f_i(\underset{\sim}{x},\nu) \odot h_i(\underset{\sim}{x},\nu) \qquad (1.99)$$

where

$$f_i(\underset{\sim}{x},\nu) = \langle |f_c(\underset{\sim}{x},\nu)|^2 \rangle$$

$$h_i(\underset{\sim}{x},\nu) = |h_c(\underset{\sim}{x},\nu)|^2, \qquad (1.100)$$

for an incoherent source. Equation (1.99) describes incoherent imaging, in which field intensities add linearly. The incoherent object $f_i(\underset{\sim}{x},\nu)$ is the spatial distribution of intensity in the object plane. The incoherent point spread function $h_i(\underset{\sim}{x},\nu)$ is equal to the squared modulus of the coherent psf.

The spatial Fourier domain representation of (1.99) is

$$S_i(\underset{\sim}{u},\nu) = F_i(\underset{\sim}{u},\nu) \, H_i(\underset{\sim}{u},\nu). \qquad (1.101)$$

The autocorrelation theorem for the Fourier transform shows that the incoherent transfer function is the autocorrelation of the coherent transfer function,

$$H_i(\underset{\sim}{u},\nu) = \int H_c^*(\underset{\sim}{u}',\nu) \, H_c(\underset{\sim}{u}+\underset{\sim}{u}',\nu) \, .d\underset{\sim}{u}'$$

$$= \mathbf{A}[H_c(\underset{\sim}{u},\nu)]$$

$$= \mathbf{A}[P(\lambda z_2 u)] \qquad (1.102)$$

where the boldface $\mathbf{A}[\cdot]$ stands for the autocorrelation operator.

(1.101) is usually written in the normalised form

$$\hat{S}_i(\underset{\sim}{u},\nu) = \frac{S_i(\underset{\sim}{u},\nu)}{S_i(\underset{\sim}{0},\nu)}$$

$$= \frac{F_i(\underset{\sim}{u},\nu)}{F_i(\underset{\sim}{0},\nu)} \frac{H_i(\underset{\sim}{u},\nu)}{H_i(\underset{\sim}{0},\nu)}$$

$$= \hat{F}_i(\underset{\sim}{u},\nu)\, \hat{H}_i(\underset{\sim}{u},\nu). \qquad (1.103)$$

$\hat{S}_i$, $\hat{F}_i$, and $\hat{H}_i$ are all equal to unity at the origin of $\underset{\sim}{u}$-space. The normalised transfer function $\hat{H}_i(\underset{\sim}{u},\nu)$ is known as the optical transfer function or OTF of the imaging instrument, and its magnitude $|\hat{H}_i(\underset{\sim}{u},\nu)|$ is called the modulation transfer function or MTF [Goodman 1968 φ6.3]. The normalised object visibility $\hat{F}_i(\underset{\sim}{u},\nu)$ is equal to the normalised mutual coherence (φ1.4.2) of the field in the aperture plane of the imaging instrument at any two points separated by $\underset{\sim}{\xi} = \lambda z_2 \underset{\sim}{u}$, according to the van Cittert-Zernike theorem [Born and Wolf 1975 φ10.4.2; cf. Leader 1978].

A geometrical interpretation of the OTF $\hat{H}_i$ for a diffraction-limited instrument is presented in Figure 1.8. Figure 1.8a shows the coherent transfer function, or pupil function, $H_c(\underset{\sim}{u}',\nu)$, whose area corresponds to $H_i(\underset{\sim}{0},\nu)$. Figure 1.8b shows two versions of $H_c(\underset{\sim}{u}',\nu)$ superposed with an offset of $\underset{\sim}{u}$. The area of

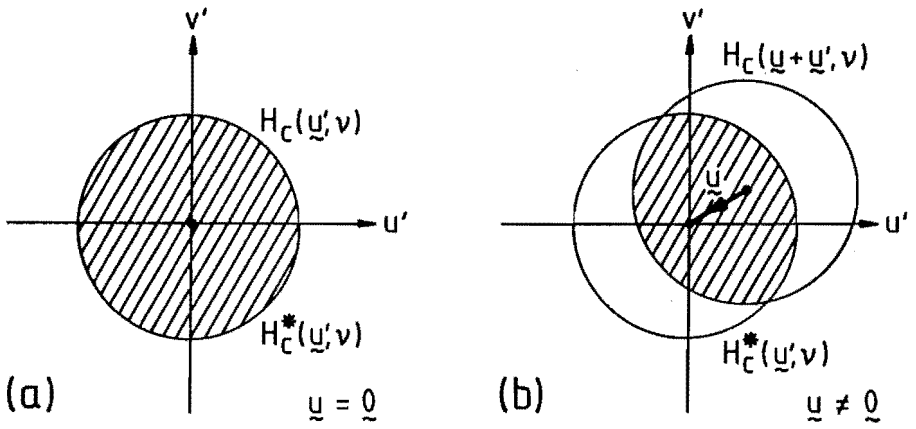


Figure 1.8: A geometrical interpretation of the OTF of a diffraction-limited imaging instrument, in terms of the overlap of two versions of the pupil function displaced relative to each other by the spatial frequency vector $\underset{\sim}{u}$. The OTF is proportional to the shaded area of overlap.
(a) $\underset{\sim}{u} = \underset{\sim}{0}$. (b) $\underset{\sim}{u} \neq \underset{\sim}{0}$.

overlap corresponds to $H_i(\underset{\sim}{u},\nu)$. The ratio of the overlap area to the total area of the pupil function gives the OTF $\hat{H}_i(\underset{\sim}{u},\nu)$. Evidently $\hat{H}_i$ has a maximum value of unity for $\underset{\sim}{u}$ = zero, decreases as $|\underset{\sim}{u}|$ increases from zero, and is zero for all $|\underset{\sim}{u}|$ greater than the maximum diameter $\frac{D}{\lambda z_2}$ of $H_c(\underset{\sim}{u},\nu)$ (where D is the maximum diameter of the instrument's aperture; cf. (1.96)). The frequency $\frac{D}{\lambda z_2}$ is sometimes called the diffraction limit.

The spatial frequency $\underset{\sim}{u}$ in an incoherent imaging system

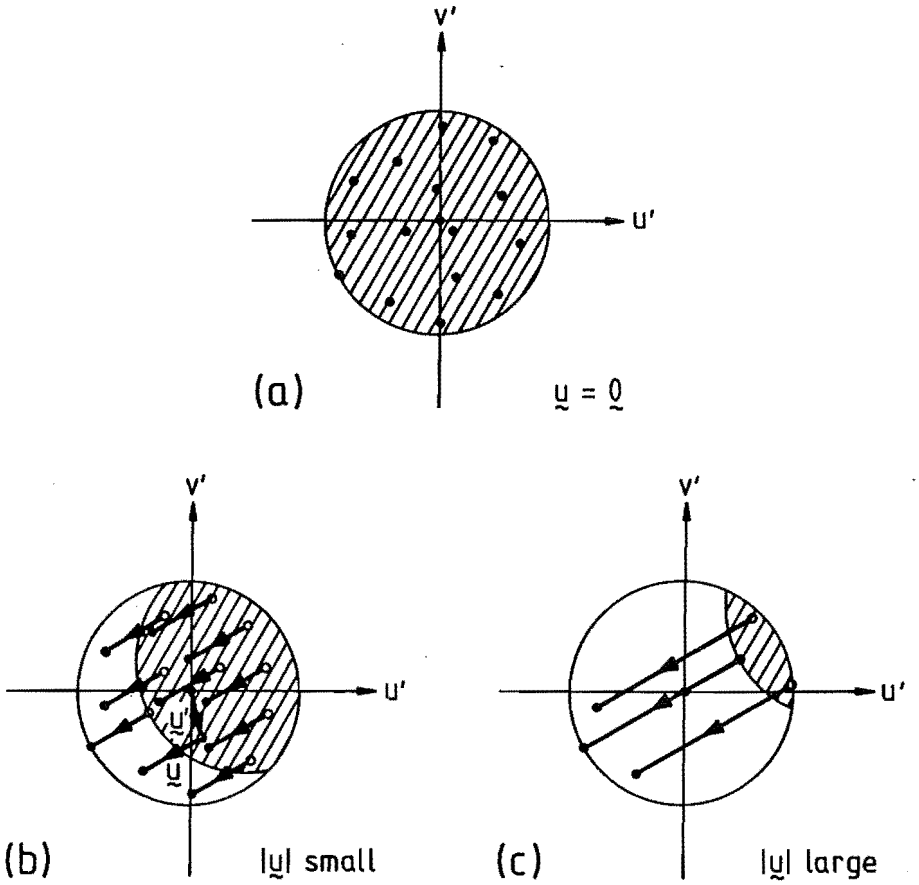


Figure 1.9: Another geometrical interpretation of the OTF of a diffraction-limited imaging instrument, in terms of displacement vectors $\underset{\sim}{u}$ within the aperture. The OTF is proportional to the number of displacement vectors whose ends both fit within the aperture.

(a) $\underset{\sim}{u} = \underset{\sim}{0}$: the maximum number of vectors fit within the aperture.

(b) $|\underset{\sim}{u}|$ small: many vectors fit within the aperture.

(c) $|\underset{\sim}{u}|$ large: few vectors fit within the aperture.

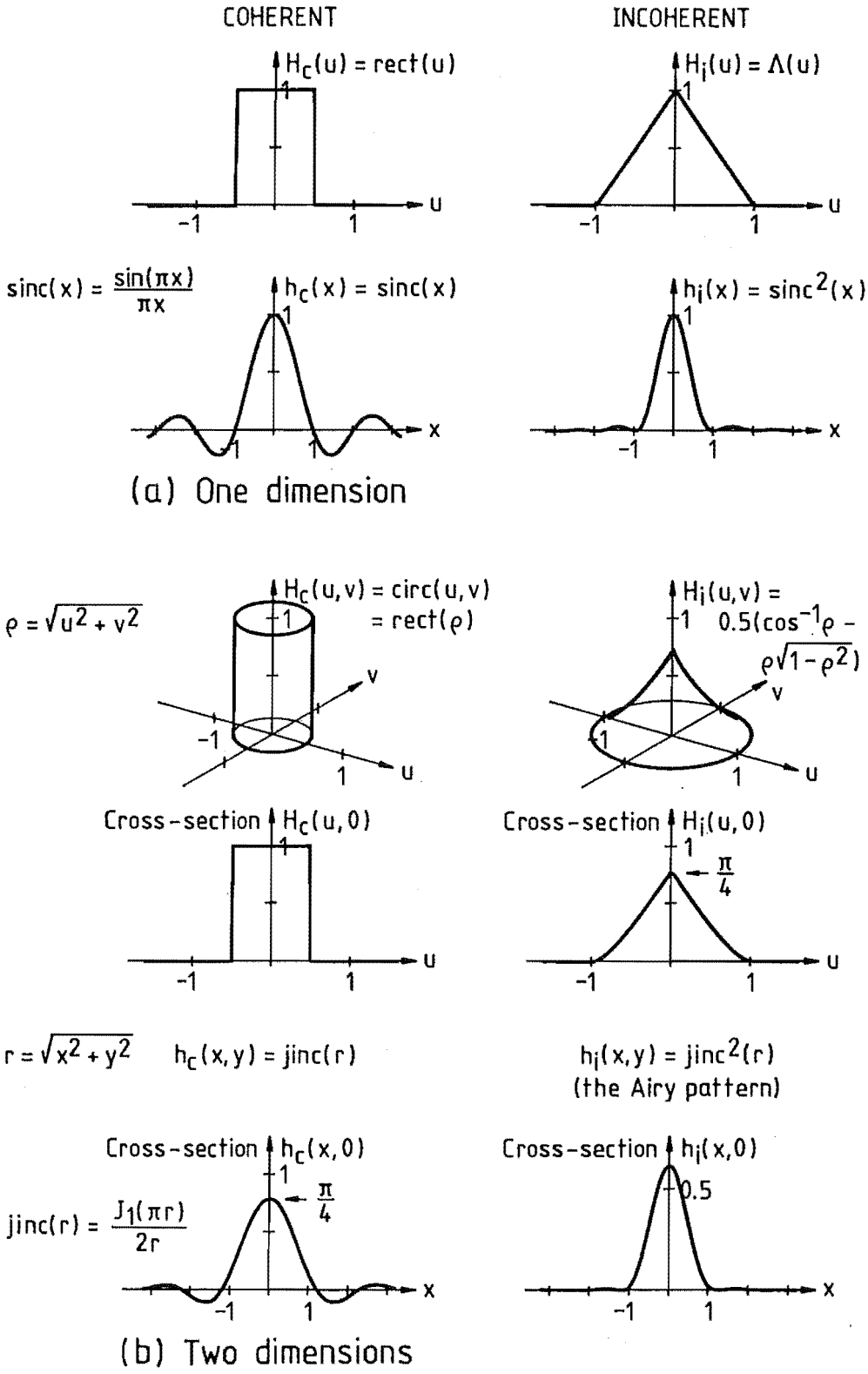


Figure 1.10: Diffraction-limited transfer functions and point spread functions for (a) a one-dimensional linear aperture and (b) a two-dimensional circular aperture.

represents a displacement vector in the aperture plane, whose absolute position is unimportant. A second geometrical interpretation of the OTF is presented in Figure 1.9, in which the OTF is represented by the number of displacement vectors $\underset{\sim}{u}$ whose ends fit within the aperture of the imaging instrument, relative to the number of points (displacement vectors of zero length) within the aperture. These numbers are proportional to the shaded areas in Figure 1.9, which are equivalent to the overlap areas in Figure 1.8.

Figure 1.10 shows examples of diffraction-limited transfer functions and point spread functions in one and two dimensions. The diffraction limit $|\underset{\sim}{u}| = \frac{D}{\lambda z_2}$ (cf. 1.96) has been normalised to unity in these examples. The seeming anomaly of the incoherent imaging system having a spatial frequency cutoff twice that of a coherent imaging system is discussed by Goodman [1968 ϕ6.5].

1.6.5 ABERRATIONS

A diffraction-limited imaging system, when forming an image of a point object, produces from its aperture a perfect spherical wave converging towards the ideal image point. This perfection is unattainable in practice, since any real-world imaging system in a similar circumstance produces a converging wave with deviations from perfect sphericity. These deviations are called aberrations [Born and Wolf 1975 ϕϕ5,9]. Many aberrations can be usefully represented as distortions in the phase of the converging wave. These are conveniently described by replacing the pupil function $P(\underset{\sim}{\xi})$ (1.90) with a complex generalised pupil function

$$Q(\underset{\sim}{\xi}) = P(\underset{\sim}{\xi})\, e^{j\phi(\underset{\sim}{\xi})} \qquad (1.104)$$

where $\phi(\underset{\sim}{\xi})$ represents the phase deviations across the aperture, which might, for example, arise from optical path length differences due to figuring errors in the lens or mirror surfaces of an optical instrument.

The effect of phase fluctuations in $Q(\underset{\sim}{\xi})$ on the OTF $\hat{H}_i(\underset{\sim}{u},\nu)$ is to make it complex-valued and, in general, to reduce its magnitude for

values of $\underset{\sim}{u}$ other than zero. Referring to Figure 1.9, the calculation of the OTF when the pupil function is complex involves multiplying the value of H_c at one end of the displacement vector $\underset{\sim}{u}$ by the complex conjugate of H_c at the other end, and summing these complex products for all positions $\underset{\sim}{u}'$ of the displacement vector (cf. the integral in (1.102)). When the phase of H_c is different at each end of $\underset{\sim}{u}$, the complex product has nonzero phase which can lead to partial or complete cancellation when it is summed over the whole aperture, thereby depressing the MTF $|\hat{H}_i|$. Thus, aberrations in an incoherent imaging system tend to reduce the contrast of detail in the image, and may even reverse the contrast [Goodman 1968 ∅6.4].

When the phase of H_c is uncorrelated or random across the length of $\underset{\sim}{u}$, the summed product referred to above has a random phase and a random magnitude. The random magnitude tends towards zero as the area of summation (Figures 1.8, 1.9) increases. For a particular displacement $\underset{\sim}{u}$, the number of displacement vectors that fit within the aperture (i.e. the overlap area in Figure 1.8), whose increase leads to a reduction in the MTF, has been called the redundancy of the aperture [Russell and Goodman 1971, Rhodes and Goodman 1973, Roddier 1981 ∅5.2, Greenaway 1982; cf. Arsac 1955].

While the imagery of an ideal diffraction-limited imaging instrument (1.93, 1.99) is isoplanatic, the effects of aberrations in a practical instrument are generally different in different parts of the image plane, i.e. real-world instruments are nonisoplanatic. The notion of effective isoplanatism of the instrument in isoplanatic patches of limited size (∅1.2.1) is useful here, particularly in circumstances such as optical astronomy in which the presence of an inhomogeneous medium between the object and the imaging instrument produces nonisoplanatic effects (∅2.6) much greater than those of the instrument itself.

1.6.6 RESOLUTION

The resolution of an imaging instrument is the size of the finest discernible detail it can render in its image. This is closely connected with the extent and shape of the imaging instrument's point

spread function. Quantitative definitions of resolution have to be formulated arbitrarily because the fineness of discernible detail depends on the form of the psf and also on the nature of the object.

The Rayleigh criterion, which was originally introduced to describe the resolving power of spectrometers [Born and Wolf 1975 ϕ7.6.3], has been widely used to describe the resolution of two point objects by an instrument with an unobscured circular aperture. The incoherent point spread function for these instruments is the Airy pattern (Figure 1.10), which consists of a bright central lobe (the Airy disc) surrounded by alternating dark and bright rings (the Airy rings). For an instrument with an aperture of diameter D, the radius of the first dark ring is given by

$$r_0 \approx \frac{1.22\lambda z_2}{D} \tag{1.105a}$$

in terms of distance in the image plane, which referred to angular position in the object in radian units (replacing $-\ell$ by z_2 in (1.88)) becomes

$$\alpha_0 \approx \frac{1.22\lambda}{D} \ . \tag{1.105b}$$

The Rayleigh criterion holds that two point objects are barely resolved when the centre of the Airy disc from one object coincides with the first dark ring from the other, i.e. they are resolved when their angular separation is equal to or greater than α_0. The combined image in the barely resolved case has two intensity maxima with a slight dip between them.

The half-intensity diameter of the Airy disc provides another reasonable measure of resolution for a circular aperture. It is approximately equal to $1.02\ \frac{\lambda}{D}$ in angular measure (cf. Figure 1.10), which is similar to the Rayleigh limit (1.105). The half-intensity diameter is analogous to the half-power beamwidth of an antenna or the half-power bandwidth of an electrical filter.

The integral over the OTF of an incoherent imaging system can be used as a measure of resolution [Fried 1966]. This integral,

$$R = \int \hat{H}_i(\underset{\sim}{u},\nu)\ .d\underset{\sim}{u} = \frac{\int H_i(\underset{\sim}{u},\nu)\ .d\underset{\sim}{u}}{H_i(\underset{\sim}{0},\nu)}\ , \tag{1.106}$$

expresses the equivalent area [cf. Bracewell 1978 p.151] of the transfer function $H_i(\underset{\sim}{u},\nu)$. The equivalent area of the psf $h_i(\underset{\sim}{x},\nu)$, which is the reciprocal of this [cf. Bracewell 1978 p.151], provides an inverse measure of resolution (the larger it is, the poorer is the resolution). For a circular aperture of diameter D, the equivalent area of the psf $h_i(\underset{\sim}{x},\nu)$ is given in terms of linear measure in the image plane by

$$\frac{1}{R_r} = \frac{4}{\pi}\left(\frac{\lambda z_2}{D}\right)^2\ . \tag{1.107a}$$

In terms of angular measure in the object this becomes

$$\frac{1}{R_\alpha} = \frac{4}{\pi}\left(\frac{\lambda}{D}\right)^2 \tag{1.107b}$$

[Roddier 1981 ϕ4.3].

1.7 SUMMARY

In many situations of practical interest, wave propagation is described to a good accuracy by a linear, scalar model. From this description comes the branch of optical theory known as Fourier Optics [Goodman 1968]. The Fourier transform, which is a useful tool in studying the behaviour of any linear system, appears explicitly in Fourier optics in the relation between the complex amplitude of a source and its far field radiation pattern, the intensity of a spatially incoherent source and the mutual coherence of its radiation in the far field, and the complex amplitudes of the fields in the front and back focal planes of an ideal thin lens. The Fourier transform expresses a quantity as the sum of sinusoidal components of different frequencies. The spatial Fourier transform of a two-dimensional source has the form of an angular spectrum of plane waves propagating in three-dimensional space, which is physically analogous to the far field radiation of the source. The response of an imaging system to different spatial frequencies is described by its

transfer function. Spatial frequency space and the transfer function provide a convenient description of the effects of diffraction limiting and of some types of aberrations on the performance of imaging systems.

CHAPTER 2

OPTICAL ASTRONOMICAL IMAGING

2.1 INTRODUCTION

In conventional optical astronomical imaging, celestial objects are viewed through the Earth's atmosphere with a ground-based telescope. The telescope aperture is usually circular with only small obstructions and may be several metres in diameter in large instruments [Barlow 1975]. The telescope is aimed at a celestial object and guided to follow the object's apparent motion across the sky. An image is formed by a detector such as photographic film or a photoelectric device [Eccles et al 1983] which integrates the intensity of light at the telescope's image plane over a certain exposure time. For the long exposure times typical of conventional astronomical imaging the spatial resolution of the image is limited not by the optical performance of the telescope but by the optical properties of the Earth's atmosphere, unless the aperture is very small (say ≤ 10cm diameter) [Barlow 1975 ϕ4.17]. Randomly fluctuating variations of refractive index in the atmosphere perturb the wavefronts of light travelling from the celestial objects to the telescope, distorting the image. Images formed with exposure times any longer than about one second are blurred so that an unresolved star (i.e. one which would not be resolved by the telescope without atmospheric effects; this is true for all stars except the Sun and a number of red giants in our galaxy) appears as a fuzzy patch called the seeing disc. The angular size of the seeing disc depends on the quality of the seeing, i.e. the severity of the atmospheric perturbations to the incoming light. For typical "good" seeing the diameter of the seeing disc is roughly one arc second.

The size of the seeing disc imposes a limit on the spatial resolution of conventional optical astronomical images. Its size (for the typical value quoted above) corresponds to the size of the diffraction-limited image of an unresolved star produced by a circular aperture of roughly 10cm diameter. The potential (diffraction-

limited) resolution of the World's largest optical telescopes is far better than this, e.g. about 0.02 arc seconds for a 5m diameter aperture, which is a factor of 50 or more better than the resolution obtained in practice. Large apertures collect more light, allowing fainter objects to be imaged, but do not improve the resolution. This is the seeing problem in optical astronomy.

In this chapter the optical astronomical imaging system and the effects of the Earth's atmosphere on optical astronomical imaging are reviewed. A general description of a typical optical astronomical imaging system is given in Section 2.2. The statistics of the optical field at the aperture of the telescope are described in Section 2.3. Section 2.4 considers the long time average of the image plane intensity, which is the conventional long exposure image. Section 2.5 considers the "instantaneous" image plane intensity, which is recorded in short exposure images with exposures much shorter than one second. Nonisoplanatism due to the atmosphere is treated in Section 2.6. In Section 2.7 a simple "phase screen" model for the atmosphere is described. This model is the basis of the algorithms described in Chapter 4 for simulating short exposure images with a digital computer. Section 2.8 briefly describes some of the techniques, other than speckle processing (see Chapter 3), which have been developed to obtain high spatial resolution despite atmospheric seeing.

Published review articles on the effects of the atmosphere on optical propagation and astronomy have been written by Strohbehn [1971] and Roddier [1981] with a theoretical approach and by Woolf [1982] from the viewpoint of the astronomical observer. Ricort and Aime [1979 ϕIV.A] provide an excellent concise summary.

2.2 OPTICAL ASTRONOMICAL IMAGING SYSTEM

The imaging system of optical astronomy is shown schematically in Figure 2.1. The sources are celestial bodies such as stars, galaxies, nebulae, and planets and moons, which radiate or reflect light. The sources exist at various distances from the telescope, but the distances are all so great as to be effectively infinite as far as telescope imagery is concerned (Figure 1.7c). Only the angular

coordinates of the sources in the heavens are directly inferable from the telescope image, so it is sensible to (conceptually) replace the actual celestial sources by two-dimensional equivalent sources lying on a celestial sphere of large radius centred on the telescope aperture.

The imaging instrument is an optical telescope [Barlow 1975], which forms an image of a region of the celestial sphere in its focal plane. The angular extent of the region imaged by the telescope is very small in high resolution optical astronomy: even the wide-angle Schmidt "cameras" used in star mapping rarely have fields of view exceeding 10° [Barlow 1975 §4.4]. Angular coordinates $\underset{\sim}{\alpha}$ (in radians) on the celestial sphere are related to linear position $\underset{\sim}{x}$ in the image plane by the telescope's focal length ℓ,

$$\underset{\sim}{x} \approx -\ell\underset{\sim}{\alpha} \tag{2.1}$$

for small $|\underset{\sim}{\alpha}|$ (cf. 1.88).

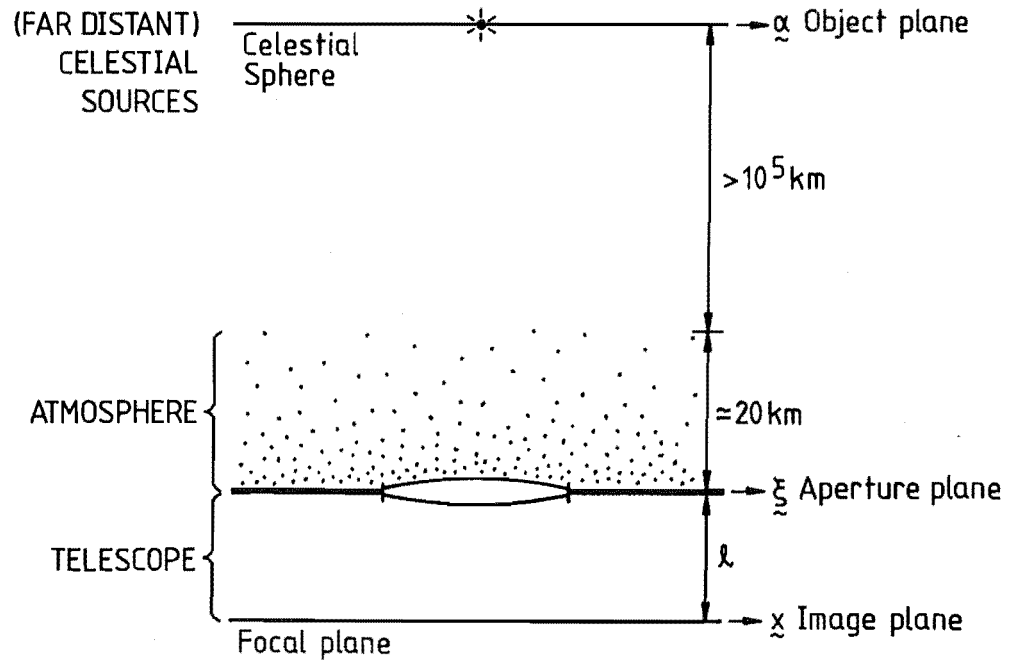


Figure 2.1: Schematic diagram of the imaging system of optical astronomy. Light from celestial sources (e.g. stars) passes through the atmosphere to a telescope at the Earth's surface. Images are formed in the focal plane of the telescope.

Celestial sources are spatially incoherent (§1.4.2.2) at optical frequencies. The equivalent sources on the celestial sphere also appear spatially incoherent, because the radius of the celestial sphere is exceedingly large compared to the dimensions of the telescope aperture [Bates 1982b §2.3]. The object in optical astronomy is therefore taken to be the two-dimensional distribution $f(\underset{\sim}{\alpha})$ of intensity on the celestial sphere. The image is the intensity $s(\underset{\sim}{x})$ of the field in the image plane of the telescope. For convenience, the object and the image are often expressed as functions of the same spatial position vector, either $\underset{\sim}{\alpha}$ or $\underset{\sim}{x}$, using (2.1). In this thesis they are both written as functions of linear position $\underset{\sim}{x}$ in the image plane of the telescope. The "object" $f(\underset{\sim}{x})$ has physical significance as the image that would be formed by an ideal instrument with no aberrations and no diffraction limiting (§§1.6.2, 1.6.5).

The Earth's atmosphere exists as an inhomogeneous medium interposed between the celestial sources and the telescope. The inhomogeneities affecting the propagation of light are variations of refractive index caused principally by variations in temperature in the atmosphere [Roddier 1981]. The atmosphere is in continual, turbulent motion, so the inhomogeneities vary randomly in space and time. Causes of turbulence in the atmosphere include velocity gradients in the boundary layer with the ground, thermal plumes, wake turbulence from mountains, and wind shears at high altitudes [Woolf 1982, Roddier 1981]. An important source of image degradation is air turbulence and heating associated with the telescope dome and the telescope itself [Woolf 1982 §5]. The severity of the effects of turbulence on optical propagation is expressed by the refractive index structure constant C_N^2 [Roddier 1981 §2.3]. Typical averaged profiles of the variation of C_N^2 with height above the ground are shown in Figure 2.2 (from Roddier [1981]). Solar heating of the ground causes significant diurnal variation at low altitudes. There is a peak in $C_N^2(h)$ in the tropopause at roughly 10km altitude, caused by strong wind shears that frequently occur there [Hufnagel 1974, Barletti et al 1977]. Below about 4km altitude, disturbances to the atmosphere from terrain irregularities such as mountains cause significant variation in $C_N^2(h)$ from site to site [Barletti et al 1976, Barletti et al 1977].

Since the atmospheric inhomogeneities are random, their effects on light propagation through the atmosphere to the telescope are best described by statistical measures. A convenient approach, which is used in the following sections of this chapter, is to consider the field at the aperture plane (or pupil plane) of the telescope due to a single quasi-monochromatic point source on the celestial sphere. This is equivalent to characterising the random point spread function (ϕ1.2.1) of the atmosphere. An unresolved source is one whose apparent angular size on the celestial sphere is sufficiently less than the diffraction-limited resolution of the telescope (ϕ1.6.6) that its image is indistinguishable from that of a point source.

Radiation from a celestial point source has effectively planar wavefronts at the top of the atmosphere, because the celestial sphere is very distant from the Earth. The thickness of the atmosphere is miniscule compared with this distance (Figure 2.1). The planar wavefronts are perturbed as the radiation propagates downward through the atmosphere to the telescope. The field at the aperture plane of the telescope exhibits random spatial and temporal variations of magnitude and phase due to the atmospheric perturbations. The magnitude variations are known as scintillation [Roddier 1981 ϕ8], and can be pictured as coming from the varying areal density at the

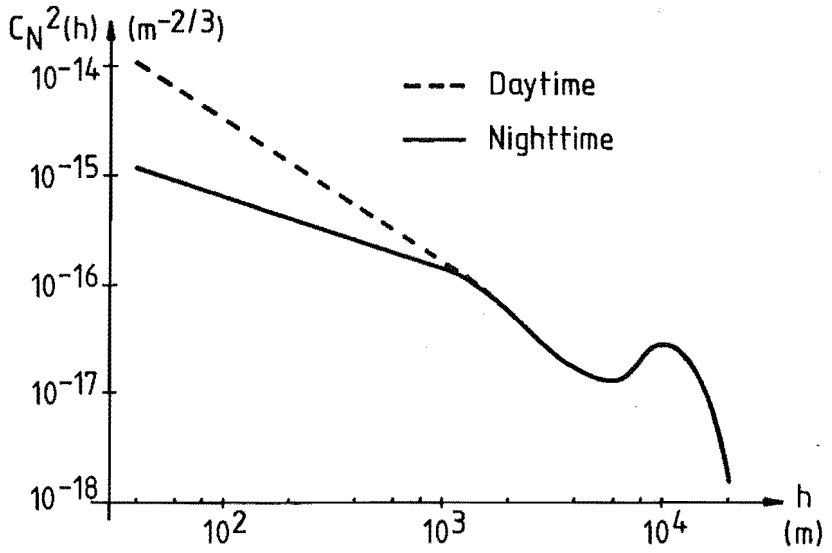


Figure 2.2: Typical averaged profile of the refractive index structure constant $C_N^2(h)$ (after Roddier [1981]).

telescope aperture of light rays bent by refraction in the high regions of the atmosphere. Scintillation is manifest to the naked eye as the twinkling, or time-varying brightness, of the stars. The phase variations, which span several times the principal range of 2π radians at optical frequencies [Breckinridge 1976, Scaddan and Dainty 1977], come from the variations of integrated refractive index along the rays through the atmosphere to each part of the telescope pupil.

The magnitude and phase fluctuations of the aperture plane field can be lumped together when considering the effects of the atmosphere on the telescope image [Fried 1966, Roddier 1981]. The contribution of the magnitude fluctuations to image degradation is generally much less than that of the phase fluctuations [Young 1974, Roddier 1981 §7.3], and can often be omitted in analyses of imaging through the atmosphere [e.g. Roddier 1981 §§4,5]. This approximation is called the near field approximation [Roddier 1981 §7.3]. It is equivalent to assuming that the atmosphere perturbs the speed but not the direction of light rays passing through the atmosphere to the telescope. The aperture plane field from a quasi-monochromatic point source of mean wavelength $\bar{\lambda}$ and unit strength is described under the near field approximation by

$$\psi(\underline{\xi},\bar{\lambda},t) = e^{j\phi(\underline{\xi},\bar{\lambda},t)} \tag{2.2}$$

where $\phi(\underline{\xi},\bar{\lambda},t)$ is the random phase fluctuation across the telescope aperture. This is a phasor description (§1.2.2), with the t-dependence describing the effects of the atmosphere's motion.

In considering time-averaged quantities, such as spatial coherence and intensity of the aperture plane and image plane fields, and the short and long exposure images, several different time intervals need to be defined. Following Bates [1982b §2.4], the temporal fluctuations of the atmosphere are characterised here by the short-interval τ_s and the redistribution time τ_r. τ_s is the longest interval over which the effects of the atmosphere can be considered to be unchanging or "frozen", so that $\psi(\underline{\xi},\bar{\lambda},t)$ is effectively constant during this interval. τ_r is the shortest interval between instants at which the states of the atmosphere affecting $\psi(\underline{\xi},\bar{\lambda},t)$ are statistically independent. Typical values quoted for these intervals

are of the order of 10ms for τ_s and 500ms for τ_r [Bates 1982b ϕ2.4] (cf. ϕ2.5).

"Instantaneous" values of the averaged field quantities are evaluated over an interval long enough to allow the spatial incoherence of the source to become established (ϕ1.4.2). It is safe to assume that this interval is much less than τ_s [Bates 1982b ϕ2.3]. A short time average is taken over an interval that is not shorter than the "instantaneous" interval and not longer than τ_s. A long time average is taken over an interval equal to at least several redistribution times. The atmospheric fluctuations are assumed to be statistically stationary during a long time average. Short and long time averages are denoted in this chapter, where necessary, by

$$\langle\cdot\rangle_{ST}, \quad \langle\cdot\rangle_{LT}$$

respectively. In the following chapters, however, long time averages are treated as averages over an ensemble of short exposure images (ϕ2.5) taken at different instants (see ϕ3.2).

2.3 STATISTICS OF THE APERTURE PLANE FIELD AMPLITUDE

2.3.1 FIRST-ORDER STATISTICS

The first-order statistics of the complex amplitude of the aperture plane field due to a quasi-monochromatic point source are described with reasonable accuracy, for astronomical imaging at small or moderate zenith angles, by the log-normal model [Korff 1973, Roddier 1981 ϕϕ7,8]. This model describes the log-modulus and the phase of the field as independent normally distributed random variates [Tatarski 1961]. The log-normal model is a consequence of the Central Limit theorem in forward scattering or line-of-sight propagation, where the perturbations in the aperture plane field are a product of contributions from the medium through which the radiation has passed, and the Rytov transformation (ϕ1.5.4.2) is appropriate [Ishimaru 1978 ϕ6.6].

An alternative model is the normal or complex Gaussian model, which describes the real and the imaginary parts of the aperture-plane field as independent normally distributed random variates with zero means. This is considered to be a reasonable approximation when the seeing is poor, and has been used in place of the log-normal model for mathematical convenience [e.g. Korff et al 1972, Dainty 1973, Dainty 1974, Roddier et al 1982]. Experimental measurements favour the log-normal model over the complex Gaussian model at visible wavelengths [Aime et al 1979] and in the infrared [Chelli et al 1979].

The angular extent of a celestial source has a marked effect on the scintillation in the aperture plane field. This is why the stars twinkle more than the planets do [Tatarski 1961 ϕ13] when observed with the naked eye. Information about the angular structure of a source can be inferred from measurements of the scintillation [Little and Hewish 1966]. This has been termed scintillation interferometry [Bates and Gough 1975, Bates 1982b ϕ11.2].

2.3.2 SECOND-ORDER STATISTICS

The mutual coherence function (ϕ1.4.2) is an expression of the second-order statistics of field amplitude. In an incoherent imaging system (ϕ1.6.4) the second-order statistics of the aperture plane field are of crucial concern because the mutual coherence at the aperture plane is the spatial Fourier transform of the incoherent image (cf. ϕ1.6.4).

Roddier [1981] considers the case of a monochromatic plane wave propagating downward through the atmosphere to the telescope aperture. The atmosphere is modelled by a continuum of non-absorbing horizontal "phase screens". The refractive index fluctuations in each phase screen are assumed to be accurately described by the propagation theory presented by Tatarski [1961], and are characterised by the refractive index structure constant $C_N^2(h)$ as a function of height h. On integrating the effects of each phase screen through the whole height of the atmosphere, the mutual coherence function at the telescope aperture plane is found, under the near-field approximation (ϕ2.2), to be [Roddier 1981 ϕ3.2]

$$B_{LT}(\underset{\sim}{\xi},\lambda) = \langle \psi^*(\underset{\sim}{\xi}',\lambda,t)\,\psi(\underset{\sim}{\xi}+\underset{\sim}{\xi}',\lambda,t)\rangle_{LT}$$

$$= e^{-\frac{1}{2}\left[2.91k^2(\cos\gamma)^{-1}|\underset{\sim}{\xi}|^{\frac{5}{3}}\int C_N^2(h).dh\right]} \qquad (2.3)$$

where $\underset{\sim}{\xi}$ is the spatial coordinate in the aperture plane, k is the wave number of the light (1.43), and γ is the magnitude of angle from the zenith. Roddier [1981 φ4.5] summarises experimental measurements of $B_{LT}(\underset{\sim}{\xi},\lambda)$, which in general agree well with (2.3).

2.4 LONG EXPOSURE IMAGE: THE SEEING DISC

The long exposure image is the long time average of the intensity of the image plane field. The seeing disc is the long exposure image of an unresolved celestial object. The seeing disc is equivalent to the point spread function of the telescope and atmosphere for long exposure imaging.

If the telescope and atmosphere are assumed to be isoplanatic over the field of view of the telescope for long exposures then the spatial Fourier transform of the seeing disc can be defined as the long exposure transfer function $H_{LE}(\underset{\sim}{u},\nu)$ of the telescope/atmosphere system. Under the more restrictive assumption of isoplanatism in the "instantaneous" image, the form of $H_{LE}(u,\nu)$ may be derived as follows.

The "instantaneous" coherent transfer function (cf. φ1.6.3) $H_c(\underset{\sim}{u},\nu,t)$ of the telescope and atmosphere is equal to the product of the telescope's pupil function $Q(\underset{\sim}{\xi})$ and the field $\psi(\underset{\sim}{\xi},\lambda,t)$ at the pupil of the telescope due to a point object of unit strength at the origin of object space on the celestial sphere, i.e.

$$H_c(\underset{\sim}{u},\nu,t) = Q(\lambda\ell\underset{\sim}{u})\;\psi(\lambda\ell\underset{\sim}{u},\lambda,t)\;. \qquad (2.4)$$

"Unit strength" here means that, in the absence of atmospheric turbulence, $\psi(\underset{\sim}{\xi},\lambda,t) = 1$ (cf. 2.2). The "instantaneous" incoherent transfer function $H(\underset{\sim}{u},\nu,t)$ is equal to the autocorrelation of (2.4) (cf. 1.102),

$$H(\underset{\sim}{u},\nu,t) = \mathbf{A}[H_c(\underset{\sim}{u},\nu,t)]$$

$$= \int Q^*(\lambda\ell\underset{\sim}{u}')\,Q(\lambda\ell(\underset{\sim}{u}+\underset{\sim}{u}'))\,\psi^*(\lambda\ell\underset{\sim}{u}',\lambda,t)\,\psi(\lambda\ell(\underset{\sim}{u}+\underset{\sim}{u}'),\lambda,t)\,.d\underset{\sim}{u}'. \tag{2.5}$$

The long exposure transfer function $H_{LE}(\underset{\sim}{u},\nu)$ is the long time average of the "instantaneous" incoherent transfer function,

$$H_{LE}(\underset{\sim}{u},\nu) = \langle H(\underset{\sim}{u},\nu,t)\rangle_{LT}$$

$$= \int Q^*(\lambda\ell\underset{\sim}{u}')\,Q(\lambda\ell(\underset{\sim}{u}+\underset{\sim}{u}'))\,\langle\psi^*(\lambda\ell\underset{\sim}{u}',\lambda,t)\,\psi(\lambda\ell(\underset{\sim}{u}+\underset{\sim}{u}'),\lambda,t)\rangle_{LT}\,.d\underset{\sim}{u}'. \tag{2.6}$$

The time-averaged quantity within the integral in (2.6) is the mutual coherence function (2.3), which is independent of $\underset{\sim}{u}'$ and so may be brought outside the integral to give

$$H_{LE} = \int Q^*(\lambda\ell\underset{\sim}{u}')Q(\lambda\ell(\underset{\sim}{u}+\underset{\sim}{u}'))\,.d\underset{\sim}{u}'\ \langle\psi^*(\lambda\ell\underset{\sim}{u}',\lambda,t)\psi(\lambda\ell(\underset{\sim}{u}+\underset{\sim}{u}'),\lambda,t)\rangle_{LT}$$

$$= H_T(\underset{\sim}{u},\nu)\,H_A(\underset{\sim}{u},\nu), \tag{2.7}$$

showing that the long exposure transfer function of the telescope plus atmosphere is the product of individual transfer functions for the telescope and the atmosphere [Hufnagel and Stanley 1964, Fried 1966, Roddier 1981]. The relations leading to (2.7) are centrally important when considering astronomical imaging and so they are collected together and restated here for reference:

$$\underset{\sim}{\xi} = \lambda\ell\underset{\sim}{u}, \tag{2.8a}$$

$$H_c(\underset{\sim}{u},\nu,t) = Q(\lambda\ell\underset{\sim}{u})\ \psi(\lambda\ell\underset{\sim}{u},\lambda,t), \tag{2.8b}$$

$$H(\underset{\sim}{u},\nu,t) = \mathbf{A}[H_c(\underset{\sim}{u},\nu,t)], \tag{2.8c}$$

$$H_T(\underset{\sim}{u},\nu) = \mathbf{A}[Q(\lambda\ell\underset{\sim}{u})], \tag{2.8d}$$

$$H_A(\underset{\sim}{u},\nu) = B_{LT}(\lambda\ell\underset{\sim}{u},\lambda)$$

$$= \langle\psi^*(\lambda\ell\underset{\sim}{u}',\lambda,t)\,\psi(\lambda\ell(\underset{\sim}{u}+\underset{\sim}{u}'),\lambda,t)\rangle_{LT}, \tag{2.8e}$$

$$H_{LE}(\underset{\sim}{u},\nu) = \langle H(\underset{\sim}{u},\nu,t)\rangle_{LT}$$

$$= H_T(\underset{\sim}{u},\nu)\,H_A(\underset{\sim}{u},\nu), \tag{2.8f}$$

$$\langle h(\underset{\sim}{x},\nu,t)\rangle_{LT} \leftrightarrow \langle H(\underset{\sim}{u},\nu,t)\rangle_{LT}, \tag{2.8g}$$

$$s_{LE}(\underset{\sim}{x},\nu) = f(\underset{\sim}{x},\nu) \circledcirc \langle h(\underset{\sim}{x},\nu,t)\rangle_{LT}. \tag{2.8h}$$

The spatial resolution of long exposure images is limited primarily by whichever of H_T or H_A is narrower in $\underset{\sim}{u}$-space, since H_{LE} is their product. If the telescope aperture is of small diameter, H_T is narrow. If the seeing is poor, H_A is narrow. Resolving powers [Fried 1966] can be defined for the telescope alone and for the atmosphere alone by substituting H_T or H_A, respectively, for the H_i in (1.106). The Fried parameter r_o [Fried 1966] is defined as the diameter of the two-dimensional circular aperture of a diffraction-limited telescope for which these resolving powers are equal, i.e.

$$\int \hat{H}_T(\underset{\sim}{u},\nu).d\underset{\sim}{u} = \int \hat{H}_A(\underset{\sim}{u},\nu).d\underset{\sim}{u} \tag{2.9}$$

with $H_T(\underset{\sim}{u},\nu)$ equal to $H_D(\underset{\sim}{u},\nu)$ as given by (2.23) and (1.90). Note that $\hat{H}_A = H_A$. From (2.9), (2.3), (1.107), and (1.43), r_o is given in MKS units by

$$r_o = \left(\frac{0.060\ \lambda^2 \cos\gamma}{\int C_N^2(h).dh} \right)^{\frac{3}{5}} \tag{2.10}$$

[Barletti et al 1977; cf. Roddier 1981 φ4.3] and, substituting this into (2.3) (cf. 2.8e), the atmospheric transfer function takes the form

$$H_A(\underset{\sim}{u},\nu) = e^{-3.44\left(\frac{\lambda \ell |\underset{\sim}{u}|}{r_o}\right)^{\frac{5}{3}}} \tag{2.11}$$

which is illustrated and compared with H_T for a large telescope in Figure 2.3. Note from (2.10) that r_o increases with wavelength [Karo and Schneiderman 1978] and decreases with increasing zenith angle. Experimental measurements of H_A are given by Roddier and Roddier [1973 Fig. 3], Roddier [1976], and Brown and Scaddan [1979]; see also Roddier [1981] φ4.5.

The Fried parameter r_o is an important measure of the seeing

because it characterises in a single parameter the effects of the atmosphere on conventional long exposure images. Like other properties of atmospheric turbulence, r_o varies from time to time and from place to place. Typical values at visible wavelengths are of the order of

$$r_o \approx 10\text{cm} \tag{2.12}$$

with variations of up to an order of magnitude each side of this [Labeyrie 1976, Woolf 1982 §4; Dainty and Scaddan 1974, Roddier 1976, Schneiderman and Karo 1978b, Dyck and Howell 1983, von der Lühe 1984]. The time scale of significant changes in r_o may be as short as a few seconds [Karo and Schneiderman 1976a].

By analogy with the diffraction limit $|\underset{\sim}{u}| = D/\lambda\ell$ of spatial frequencies accepted by an incoherent imaging system with an aperture of diameter D (§1.6.4), one may define the seeing limit as $|\underset{\sim}{u}| = r_o/\lambda\ell$ (cf. Figure 2.3). Seeing-limited resolution is poorer than diffraction-limited resolution when $D > r_o$. D may be fifty or more times r_o for large telescopes, which leads to a fifty-fold or more loss of resolution due to the seeing in long exposure images.

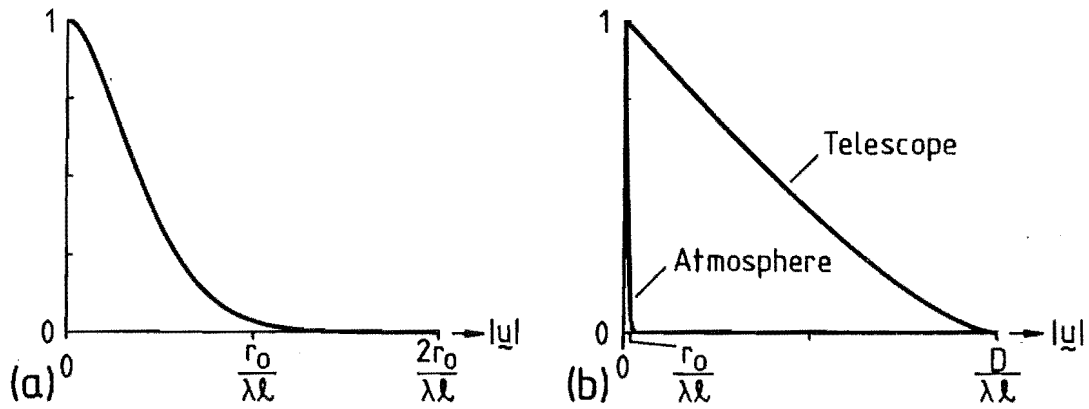


Figure 2.3: The long-exposure atmospheric transfer function $H_A(\underset{\sim}{u},\nu)$.
(a) $H_A(\underset{\sim}{u},\nu)$ from (2.11).
(b) $H_A(\underset{\sim}{u},\nu)$ (atmosphere) compared with $H_T(\underset{\sim}{u},\nu)$ (telescope) for a diffraction-limited telescope with two-dimensional circular aperture of diameter $D = 50r_o$ (e.g. 5m telescope, $r_o = 10$cm). $r_o/\lambda\ell$ is the seeing limit and $D/\lambda\ell$ is the diffraction limit.

The reciprocal of the seeing limit, λ/r_o, describes the diameter of the seeing disc in the image plane of the telescope. In angular measure on the celestial sphere (cf. 2.1) this becomes $\theta_s = \lambda/r_o$, which has been called the seeing angle [Dyck and Howell 1983]. Since r_o varies as $\lambda^{\frac{6}{5}}$ (2.10), θ_s varies as $\lambda^{-\frac{1}{5}}$ and thus the seeing slowly improves with increasing wavelength [Roddier 1981 §4.3]. The seeing angle is a measure of the resolution of long exposure images. At visible wavelengths and for typical seeing conditions (2.12) the seeing angle is approximately one arc second.

The predicted shape of the seeing disc is given by the two-dimensional Fourier transform of (2.11). Figure 2.4 (from Roddier [1981]; see also Woolf [1982]) shows a profile of this together with measured intensities from actual seeing discs [King 1971] and a Gaussian profile fitted to the central region of the seeing disc. The Gaussian profile is equivalent to replacing the exponent $\frac{5}{3}$ by $\frac{6}{3}$ in (2.11). From Figure 2.4, the measured seeing disc clearly has wider edges than the Gaussian profile. The fit between (2.11) and the

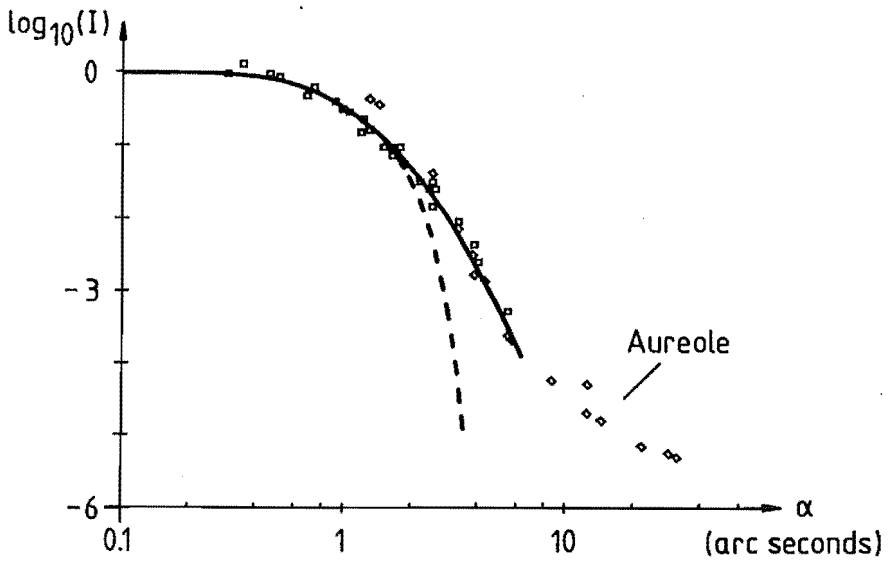


Figure 2.4: Profiles of the seeing disc (after Roddier [1981], Woolf [1982]). Vertical scale is logarithm of intensity, normalised to unity at the centre of the seeing disc. Squares and diamonds: measurements of actual seeing discs [King 1971]. Solid line: theoretical profile from the two-dimensional Fourier transform of (2.11). Dotted line: Gaussian profile.

measured seeing disc is good in both the centre and the edges. The wide aureole surrounding the measured seeing disc probably results from dust and from small imperfections in the telescope optics [Woolf 1982].

2.5 SHORT EXPOSURE IMAGE: SPECKLE

A short exposure image is an average of the image plane intensity over the short-interval τ_s, during which the state of the atmosphere is essentially constant. The short exposure image of a bright, quasi-monochromatic, unresolved source in a large telescope (of aperture diameter at least several times r_o) is a patch of light of size similar to that of the seeing disc but having a mottled or speckled appearance. The speckling arises from the interference at the image plane of light from the randomly phased field at the aperture plane. This is closely analogous to the laser speckle [Dainty 1975a] observed when light of high temporal coherence is reflected or refracted by a rough surface. Short exposure astronomical images are known as speckle images [Bates 1982b §4]. Examples of speckle images of astronomical objects have been published by (for example) Irwin [1975], Dainty [1975b Fig.7.1], and Dainty [1984 Fig.7.1]. If the short exposure imagery is assumed to be isoplanatic (this rather stringent assumption is discussed in Section 2.6), then the speckle image $s_{SE}(\underset{\sim}{x},\nu,t)$ is related to the object $f(\underset{\sim}{x},\nu)$ by (cf. 2.8)

$$\underset{\sim}{\xi} = \lambda \ell \underset{\sim}{u}, \tag{2.13a}$$

$$H_c(\underset{\sim}{u},\nu,t) = Q(\lambda\ell\underset{\sim}{u})\ \psi(\lambda\ell\underset{\sim}{u},\lambda,t), \tag{2.13b}$$

$$H(\underset{\sim}{u},\nu,t) = \mathbf{A}[H_c(\underset{\sim}{u},\nu,t)], \tag{2.13c}$$

$$h(\underset{\sim}{x},\nu,t) \leftrightarrow H(\underset{\sim}{u},\nu,t), \tag{2.13d}$$

$$s_{SE}(\underset{\sim}{x},\nu,t) = f(\underset{\sim}{x},\nu) \circledcirc h(\underset{\sim}{x},\nu,t). \tag{2.13e}$$

The "instantaneous" or short exposure transfer function $H = H(\underset{\sim}{u},\nu,t)$ cannot be separated into a product of telescope and atmosphere transfer functions, unlike the long exposure transfer function $H_{LE} = \langle H \rangle$ which can be (2.6, 2.7), because the mutual

coherence function in (2.6) evaluated as a short time average instead of a long time average (see final paragraph of ϕ2.2) is not independent of $\underset{\sim}{u}'$.

H is the autocorrelation of the random complex function H_c (2.13b). The perturbed wavefront ψ in (2.13b) is spatially correlated in pupil space with a correlation length of the order of the Fried parameter r_o. The value of H_c is correlated in $\underset{\sim}{u}$-space with a correlation length of the order of $r_o/\lambda\ell$. The expected magnitude of H is significantly greater than zero for all spatial frequencies up to the diffraction limit $|\underset{\sim}{u}| = D/\lambda\ell$, but its phase is random with essentially uniform pdf at frequencies above the seeing limit $|\underset{\sim}{u}| = r_o/\lambda\ell$ (assuming $D > r_o$). The randomness of the phase of H is the reason why the long exposure transfer function H_{LE} has essentially zero magnitude above the seeing limit (cf. Figure 2.3b, (2.8f)). For a highly redundant aperture such as a full circular aperture, the expected magnitude of the OTF $\hat{H}$ (cf. 1.103) is considerably less than the magnitude of the telescope OTF $\hat{H}_T$ for spatial frequencies above the seeing limit but appreciably below the diffraction limit. The expected squared magnitude of $\hat{H}$ above the seeing limit is proportional to the magnitude of $\hat{H}_T$, as is shown in Section 2.5.2.

The size of the speckles in a speckle image of an unresolved object is governed by the diameter of the telescope aperture, and is similar to the size of the Airy disc the aperture would form in diffraction-limited conditions. The extent of the region over which the speckles appear is governed by the atmospheric seeing (described by the Fried parameter r_o), and is similar to the size of the seeing disc. When the speckle size is not all that much smaller than the seeing disc, the speckles are less speckly than "wormy" [Bates 1982b ϕ2.1]. The number N_{sp} of speckles in the speckle image of an unresolved object is approximately equal to the number of correlation regions of the atmospheric phase distortion present within the telescope aperture [Bates 1982b ϕ4.1], and is given by

$$N_{sp} \approx \left(\frac{D}{r_o} \right)^K \qquad (2.14)$$

where K is the dimensionality of the aperture (K = 1 for a slit

aperture, K = 2 for a circular aperture) and D is the aperture's diameter. A more precise expression for the number of correlation regions in a circular telescope pupil (i.e. K = 2) is given by replacing r_o by $0.44r_o$ in (2.14) [Ricort and Aime 1979, Mariotti et al 1983, Dainty 1984 ϕ7.2.3, cf. Aime 1974].

The speckles in a speckle image change with time as the state of the atmosphere changes - they are sometimes said to "boil" [e.g. O'Donnell et al 1982]. Measurements of the temporal correlation function of speckle intensity [Parry et al 1979, Lohmann and Weigelt 1979, O'Donnell et al 1982, Ebersberger and Weigelt 1985] give temporal correlation lengths of the order of 5 to 20ms. These measurements define the length of the short-interval τ_s. The redistribution time τ_r is the interval over which the correlation becomes essentially zero. In speckle images exposed for longer than τ_s the speckles become blurred and their contrast drops. A long exposure image can be thought of as an integration of speckle images over several or many redistribution times, during which the changing speckles smear out into the featureless seeing disc. Roddier [1981] and Roddier et al [1982] point out that the effects of exposure times greater than τ_s on speckle images are formally similar to the effects of partial isoplanatism (ϕ2.6) (see Roddier [1981 ϕ6]).

The appearance of a speckle image is affected by the bandwidth of the light used to form it. A speckle image formed in wideband light does not show the same high contrast speckle as one formed with quasi-monochromatic light (1.37), because the random phase perturbations associated with a given state of the turbulent atmosphere (ϕ2.2) are wavelength dependent. However, even in a wideband speckle image, speckles are generally evident somewhere within the image, particularly near its centre [Bates and Cady 1980, Bates 1982b Fig.2.4]. To form high contrast speckle over the whole extent of the speckle image, the correlation length of the light used must be appreciably greater than the path length differences introduced by the atmospheric refractive index fluctuations and telescope aberrations [Dainty 1973]. Karo and Schneiderman [1978] have measured the averaged power spectrum (ϕ2.5.2) of speckle images formed under typical seeing conditions ($r_o \approx 10$cm) with light of various bandwidths. They find negligible loss of high spatial

frequencies from finite bandwidth for $\Delta\lambda \approx 30$nm, and significant but not large loss for $\Delta\lambda \approx 70$nm, at a centre wavelength $\bar{\lambda} = 500$nm. Wynne [1979] has described an optical system with a wavelength-dependent magnification designed to cancel the wavelength dependence of the angular scale of speckle images, and hence broaden the allowable bandwidth in speckle processing.

2.5.1 FIRST-ORDER STATISTICS OF SPECKLE IMAGES

The field at each point in the image plane is a sum of randomly-phased contributions from all parts of the telescope aperture. The phase perturbations induced by the atmosphere span several times 2π radians, as noted previously (φ2.2). When there are many correlation regions (of average diameter r_o) of the atmospheric phase perturbations within the telescope aperture (of diameter D) there are many contributions at each point in the image plane. The speckle formed in the short exposure image of a quasi-monochromatic point object is then similar to laser speckle [Dainty 1975a], and theoretical results for the first-order statistics of laser speckle [Goodman 1975] can be applied to the speckle image. From the Central Limit theorem, the first-order statistics of the complex amplitude of the image plane field are complex Gaussian. The first-order statistics of the image plane field magnitude follow the Rayleigh distribution [cf. Haykin 1978 pp.212-]. The first-order statistics of the image plane field intensity, i.e. the speckle image (φ2.5), are described by a chi-squared distribution with two degrees of freedom [Abramowitz and Stegun 1973 φ26.4], which is the negative-exponential distribution

$$p_I(i) = \Pr\{I = i\} = \frac{1}{\langle I \rangle} e^{-i/\langle I \rangle} . \qquad (2.15a)$$

$\langle I \rangle$ is the ensemble average intensity, which characterises the brightness of the object. Because of the nonzero correlation length of the aperture-plane phase fluctuations, $\langle I \rangle$ varies with position in the image plane, according to the shape of the seeing disc. The mean of the distribution (2.15a) is $\mu_I = \langle I \rangle$, and its standard deviation is $\sigma_I = \langle I \rangle$. The speckle contrast, defined as the ratio σ_I/μ_I, is equal to unity [Goodman 1975 φ2.2.3]. The cumulative distribution function

of I is given by

$$P_I(i) = \Pr\{I \leq i\} = 1 - e^{-i/\langle I \rangle} \quad . \tag{2.15b}$$

The probability density function of I normalised by its ensemble average value is

$$p_{I/\langle I \rangle}\left(\frac{i}{\langle I \rangle}\right) = e^{-i/\langle I \rangle} \quad . \tag{2.15c}$$

Equations (2.15) apply to speckle images formed with unpolarised light provided the orthogonally polarised components of each speckle image are identical in form [Goodman 1975 §2.3.3].

When the object consists of several point components, each speckle image contains several speckled contributions added incoherently. The image plane intensities add linearly for incoherent objects, so the probability density function of the speckle image can be found using characteristic functions [Bennett 1956] to be

$$p_I(i) = \sum_{n=1}^{N} \left(\frac{B_n^{N-2} \; e^{-i/B_n}}{\prod_{\substack{k=1 \\ k \neq n}}^{N} (B_k - B_n)} \right) \tag{2.16a}$$

[Goodman 1975 §2.3.2, cf. Bates 1982b §4.7] provided none of the B_n are exactly equal, where the B_n are proportional to the brightnesses of the N individual points and

$$\langle I \rangle = \sum_{n=1}^{N} B_n \quad . \tag{2.16b}$$

As the number N of point components in the object increases (in the limit, to a large extended object), the speckle image pdf tends to a normal distribution [cf. Hunt et al 1983 §5] by the Central Limit theorem, with a mean that increases as N and a standard deviation that increases only as $\sqrt{N}$. Thus the contrast of the speckles diminishes as N increases. This behaviour is markedly different from that of speckle from spatially coherent objects, e.g. in ultrasonic speckle imaging [Bates and Robinson 1981, Minard et al 1985]. In the

spatially coherent case the image plane field adds linearly by complex amplitude and so the first-order statistics of the field amplitude (and intensity) are unaffected by N except for a scaling constant [Goodman 1975 §2.3.1]. In computer simulations of coherent speckle images, Minard [1985 §6.3.2] finds that the probability density function of the image plane field magnitude is essentially identical in form for a point object and for extended objects of various shapes.

2.5.2 SECOND-ORDER STATISTICS OF SPECKLE IMAGES

The second-order statistics of speckle images can be treated by considering the averaged autocorrelation of a speckle image,

$$\langle s_{SE}(\underset{\sim}{x},\nu,t) \circledast s_{SE}(\underset{\sim}{x},\nu,t) \rangle_{LT} = \left\langle \int s^{*}_{SE}(\underset{\sim}{x}',\nu,t)\, s_{SE}(\underset{\sim}{x}+\underset{\sim}{x}',\nu,t)\, .\, d\underset{\sim}{x}' \right\rangle_{LT}, \qquad (2.17)$$

or equivalently by considering the spatial Fourier transform of the averaged autocorrelation, which is the averaged power spectrum or Wiener spectrum of the speckle image,

$$\langle |S_{SE}(\underset{\sim}{u},\nu,t)|^{2} \rangle_{LT} = \mathbf{F}[\langle s_{SE}(\underset{\sim}{x},\nu,t) \circledast s_{SE}(\underset{\sim}{x},\nu,t) \rangle_{LT}] \, . \qquad (2.18)$$

Studies of these quantities for an unresolved object $f(\underset{\sim}{x},\nu) \approx \delta(\underset{\sim}{x})f(\nu)$ give information about the point spread function $\langle \mathbf{A}[h(\underset{\sim}{x},\nu,t)] \rangle_{LT}$ and the associated transfer function $\langle |H(\underset{\sim}{u},\nu,t)|^{2} \rangle_{LT}$ of speckle interferometry (§3.3).

If the short exposure imaging process is assumed to be isoplanatic, the speckle image is related to the object by the convolution

$$s_{SE}(\underset{\sim}{x},\nu,t) = f(\underset{\sim}{x},\nu) \circledcirc h(\underset{\sim}{x},\nu,t)$$

(2.13e). The averaged power spectrum (2.18) of the speckle image is then equal to

$$\langle |S_{SE}(\underset{\sim}{u},\nu,t)|^2 \rangle_{LT} = \langle |F(\underset{\sim}{u},\nu) H(\underset{\sim}{u},\nu,t)|^2 \rangle_{LT}$$

$$= |F(\underset{\sim}{u},\nu)|^2 \langle |H(\underset{\sim}{u},\nu,t)|^2 \rangle_{LT} \qquad (2.19a)$$

$$= |F(\underset{\sim}{u},\nu)|^2 \langle |\mathbf{A}[Q(\lambda\ell\underset{\sim}{u})\,\psi(\lambda\ell\underset{\sim}{u},\lambda,t)]|^2 \rangle_{LT}$$

$$= |F(\underset{\sim}{u},\nu)|^2 \iint Q(\lambda\ell\underset{\sim}{u}')\,Q^*(\lambda\ell(\underset{\sim}{u}+\underset{\sim}{u}'))\,Q^*(\lambda\ell\underset{\sim}{u}'')\,Q(\lambda\ell(\underset{\sim}{u}+\underset{\sim}{u}''))$$
$$\langle \psi(\lambda\ell\underset{\sim}{u}',\lambda,t)\,\psi^*(\lambda\ell(\underset{\sim}{u}+\underset{\sim}{u}'),\lambda,t)\,\psi^*(\lambda\ell\underset{\sim}{u}'',\lambda,t)$$
$$\psi(\lambda\ell(\underset{\sim}{u}+\underset{\sim}{u}''),\lambda,t)\rangle_{LT}\,.d\underset{\sim}{u}'.d\underset{\sim}{u}''. \qquad (2.19b)$$

Evaluation of the fourth-order moment of ψ (the aperture plane field due to a quasi-monochromatic point source) in (2.19b) requires the assumption of a probability density function for ψ (∅2.3.1). Korff et al [1972] use an heuristic approach, modelling ψ by a collection of "cells" with independent random moduli and phases distributed across the aperture of the telescope. Dainty [1973] uses a complex Gaussian model for ψ. These approaches predict that when $D >> r_o$, $\langle |H(\underset{\sim}{u},\nu,t)|^2 \rangle_{LT}$ is of the form

$$\langle |H(\underset{\sim}{u},\nu,t)|^2 \rangle_{LT} \approx |H_{LE}(\underset{\sim}{u},\nu)|^2 + \sigma H_D(\underset{\sim}{u},\nu)$$

$$= |H_{LE}(\underset{\sim}{u},\nu)|^2 + \frac{A}{N_{sp}} H_D(\underset{\sim}{u},\nu)$$

$$= A^2 \left(|\hat{H}_{LE}(\underset{\sim}{u},\nu) + \frac{1}{N_{sp}} \hat{H}_D(\underset{\sim}{u},\nu)\right] \qquad (2.20)$$

where

$$\sigma = \int |B_{LT}(\lambda\ell\underset{\sim}{u},\lambda)|^2 . d\underset{\sim}{u} \qquad (2.21)$$

(cf. 2.8e) is the correlation area of the perturbed wavefront ψ, A is the area of the telescope aperture, N_{sp} (2.14) is the ratio of A to σ, $H_{LE}(\underset{\sim}{u},\nu)$ is the long exposure transfer function (2.8f), and

$$H_D(\underset{\sim}{u},\nu) = \mathbf{A}[\,|Q(\lambda\ell\underset{\sim}{u})|^2] \qquad (2.22)$$

is a transfer function with significant value out to the diffraction limit of the telescope. For an unapodised telescope [cf. Jacquinot and Roizen-Dossier 1964], $|Q(\lambda\ell\underset{\sim}{u})|$ is equal to either 1 or 0 (cf. 1.104, 1.90), so that (2.22) becomes

$$H_D(\underset{\sim}{u},\nu) = \mathbf{A}[P(\lambda\ell\underset{\sim}{u})] \qquad (2.23)$$

and $H_D(\underset{\sim}{u},\nu)$ is equal to the diffraction-limited incoherent transfer function of the telescope. The form of $\langle|H(\underset{\sim}{u},\nu,t)|^2\rangle$ according to (2.20) and (2.23) is the sum of the squared modulus of the long exposure transfer function (2.8f) and a component proportional to the transfer function of the telescope alone with no aberrations, provided the random aberrations of the atmosphere are worse than the fixed ones of the telescope [Dainty 1973, Dainty 1977, cf. Roddier et al 1978].

Korff [1973] uses the more realistic log-normal model for ψ to derive $\langle|H(\underset{\sim}{u},\nu,t)|^2\rangle$. His analysis predicts a low-frequency component (cf. first term on the right hand side of (2.20)) which has the form of a long exposure transfer function with image wander (§2.8) removed, and is thus somewhat wider in $\underset{\sim}{u}$-space than the first term of (2.20). The forms of $\langle|H(\underset{\sim}{u},\nu,t)|^2\rangle$ predicted by (2.20) and by Korff [1973] are nicely compared by Ricort and Aime [1976] for various values of D/r_o (see also Fried [1979]). When $D >> r_o$ the predictions of the two models are essentially identical. The crucial feature for high spatial resolution astronomy, the presence of a diffraction-limited component in $\langle|H(\underset{\sim}{u},\nu,t)|^2\rangle$, is predicted by both models. Measurements of $\langle|H(\underset{\sim}{u},\nu,t)|^2\rangle$ for unresolved stellar sources [Karo and Schneiderman 1976b, Aime et al 1979, Chelli et al 1979] agree well with Korff [1973].

The magnitude of the diffraction-limited (second) term on the right hand side of (2.20) relative to the first term, for $|\underset{\sim}{u}| = 0$, is given by the ratio

$$\frac{\sigma}{A} = \frac{1}{N_{sp}}$$

(cf. 2.14). For a large telescope with a circular aperture, $1/N_{sp}$ can be small in practice [Dainty 1975b §7.2.3], but nevertheless $\langle|H(\underset{\sim}{u},\nu,t)|^2\rangle$ has significant value nearly to the diffraction limit of the telescope [e.g. Karo and Schneiderman 1976b, Aime et al 1979] and for an aberrated telescope can even be greater than the magnitude of the transfer function $H_T(\underset{\sim}{u},\nu)$ of the telescope with no atmosphere [Dainty 1973, Dainty 1974, Karo and Schneiderman 1977; cf. Roddier et al 1978].

The second-order statistics of speckle images are described in

Fourier space by the inverse Fourier transform of (2.20),

$$\langle \mathbf{A}[h(\underset{\sim}{x},\nu,t)]\rangle_{LT} \approx \mathbf{A}[h_{LE}(\underset{\sim}{x},\nu)] + \sigma h_D(\underset{\sim}{x},\nu). \qquad (2.24)$$

For an unapodised telescope and $D >> r_o$, σ is approximately equal to the ratio of the areas in Fourier space of $|H_{LE}(\underset{\sim}{u},\nu)|^2$ and $H_D(\underset{\sim}{u},\nu)$ (cf. 2.21, 2.8, 2.23), and hence is approximately equal to the ratio of $\mathbf{A}[h_{LE}(\underset{\sim}{x},\nu)]$ and $h_D(\underset{\sim}{x},\nu)$ at the centre of image space. The central ordinates of the two terms in (2.24) are thus approximately equal to each other, allowing (2.24) to be written in normalised form as

$$\langle \hat{\mathbf{A}}[h(\underset{\sim}{x},\nu,t)]\rangle_{LT} \approx \tfrac{1}{2}\,[\hat{\mathbf{A}}[h_{LE}(\underset{\sim}{x},\nu)] + \hat{h}_D(\underset{\sim}{x},\nu)] \qquad (2.25)$$

[cf. Goodman 1975 ϕ2.5.2]. Note that $\hat{\mathbf{A}}[h_{LE}(\underset{\sim}{x},\nu)]$ and $\hat{h}_D(\underset{\sim}{x},\nu)$ are normalised to unity at the centre of image space and are not just the inverse Fourier transforms of $|\hat{H}_{LE}(\underset{\sim}{u},\nu)|^2$ and $\hat{H}_D(\underset{\sim}{u},\nu)$. According to (2.25), $\langle \mathbf{A}[h(\underset{\sim}{x},\nu,t)]\rangle_{LT}$ consists of a narrow diffraction-limited component plus a broad "fog" component. The value of the fog at the centre of image space is equal to one half of the total value of the central peak, irrespective of N_{sp}, provided $D >> r_o$.

2.6 NONISOPLANATISM

Radiation from points at different positions on the celestial sphere passes through different parts of the Earth's atmosphere to reach the aperture of the telescope, because the atmosphere occupies a volume spanning an appreciable height above the aperture. The images of the different points suffer different atmospheric distortions, making the astronomical imaging system nonisoplanatic (ϕ1.2.1). Under the near field approximation (ϕ2.2) the rays from a celestial object are assumed to be straight within the atmosphere. The part of the atmosphere affecting the image of a point source at a certain position on the celestial sphere then lies within the tunnel formed by the projection of the telescope aperture up through the atmosphere along the rays from the source (see Figure 2.5). For different source positions these tunnels enclose different volumes.

Strictly, an imaging system is either isoplanatic or

nonisoplanatic, or effectively isoplanatic or effectively nonisoplanatic. However, the descriptor of isoplanatism is the change in the form of the point spread function for different positions of the point object used to form it. If a measure of similarity between two point spread functions can be defined, it is then sensible to define a degree of isoplanatism between the two object positions which ranges from (say) 1 for perfect isoplanatism when the two point spread functions are exactly alike, to 0 for complete nonisoplanatism when the two point spread functions are completely dissimilar, according to the similarity measure used. The imaging can be said to be partially isoplanatic when the degree of isoplanatism is intermediate between 1 and 0.

The severity of the nonisoplanatism caused by the atmosphere, expressed for example as the angular diameter of the isoplanatic patch within which the degree of isoplanatism is significantly greater than zero, is in general different for different types of imaging, e.g. long or short exposures [Liu and Lohmann 1973, Wang 1975, Fried 1976,

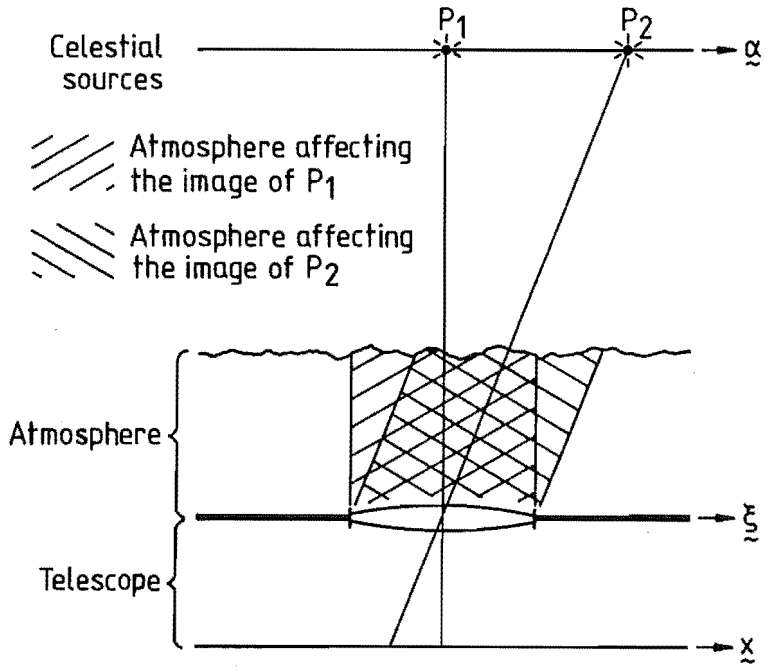


Figure 2.5: Illustrating the origin of nonisoplanatism. Radiation from different parts of the celestial sphere passes through different parts of the atmosphere to reach the telescope and hence suffers different atmospheric perturbations.

Fried 1979]. For imaging with conventional telescopes (i.e. excluding active optics), the nonisoplanatism associated with the "instantaneous" or short exposure image is of fundamental importance because all other types of image are formed by averaging or combining short exposure images in various ways.

Nonisoplanatism in the long exposure image (ϕ2.4) is due to the dependence of the form of the image on zenith angle γ (2.3, 2.8e, 2.7), if the statistical properties of the atmospheric turbulence are horizontally stationary. $\cos(\gamma)$ changes very slowly with γ when γ is small, so the effects of the variation of r_o with γ are negligible in almost all practical cases and the effects of the atmosphere on long exposure images are essentially isoplanatic [Fried 1976].

Nonisoplanatism in the short exposure image can be measured by forming a speckle image of an object comprising two unresolved sources, e.g.

$$f(\underset{\sim}{x},\nu) = \delta(\underset{\sim}{x}-\underset{\sim}{x}_1) + \delta(\underset{\sim}{x}-\underset{\sim}{x}_2), \tag{2.26}$$

$$\begin{aligned} s_{SE}(\underset{\sim}{x},\nu,t) &= \int f(\underset{\sim}{x}',\nu)\, h(\underset{\sim}{x},\underset{\sim}{x}',\nu,t).d\underset{\sim}{x}' \\ &= h(\underset{\sim}{x},\underset{\sim}{x}_1,\nu,t) + h(\underset{\sim}{x},\underset{\sim}{x}_2,\nu,t), \end{aligned} \tag{2.27}$$

where the point spread function $h(\underset{\sim}{x},\nu,t)$ in (2.13e) has been generalised to the space-variant or nonisoplanatic form $h(\underset{\sim}{x},\underset{\sim}{x}',\nu,t)$ (cf. ϕ1.2.1). If the separation $\underset{\sim}{x}_2 - \underset{\sim}{x}_1$ between the unresolved sources is sufficiently great, then the image consists of two separated speckle patterns. A quantitative measure of nonisoplanatism can be established from the cross-correlation of the speckle patterns [Lohmann and Weigelt 1979],

$$\langle h(\underset{\sim}{x},\underset{\sim}{x}_1,\nu,t) \otimes h(\underset{\sim}{x},\underset{\sim}{x}_2,\nu,t)\rangle_{LT} , \tag{2.28}$$

or equivalently from the speckle cross-spectrum (which is the Fourier transform of (2.28)) [Roddier et al 1982],

$$\langle H^*(\underset{\sim}{u},\underset{\sim}{x}_1,\nu,t) H(\underset{\sim}{u},\underset{\sim}{x}_2,\nu,t)\rangle_{LT} . \tag{2.29}$$

Expressions for the speckle cross-spectrum for a large telescope ($D >> r_o$) are derived by Roddier et al [1982], who stress that previous treatments of nonisoplanatism by Fried [1976] and Fried [1979] (see also Fried [1982], Novoseller [1985]) apply only to small apertures. Using the complex Gaussian model of aperture plane field amplitude (φ2.3.1), Roddier et al [1982] obtain the expression (expressed here in the notation of this thesis, cf. their eq. (13))

$$\langle H^*(\underset{\sim}{u},\underset{\sim}{x}_1,\nu,t)\,H(\underset{\sim}{u},\underset{\sim}{x}_2,\nu,t)\rangle_{LT} \approx |A[Q(\underset{\sim}{\xi})]|^2\,|B_{LT}(\underset{\sim}{\xi},\lambda)|^2 + \sigma(\underset{\sim}{\theta})\,Q_4(\underset{\sim}{\xi},\bar{h}\underset{\sim}{\theta}) \qquad (2.30)$$

for the speckle cross-spectrum, which applies to the geometry shown in Figure 2.6.

$$\underset{\sim}{\theta} = -(\underset{\sim}{x}_2 - \underset{\sim}{x}_1)/\ell \qquad (2.31)$$

is the angular separation of the two unresolved sources on the celestial sphere, $Q(\underset{\sim}{\xi})$ is the telescope's pupil function (1.104),

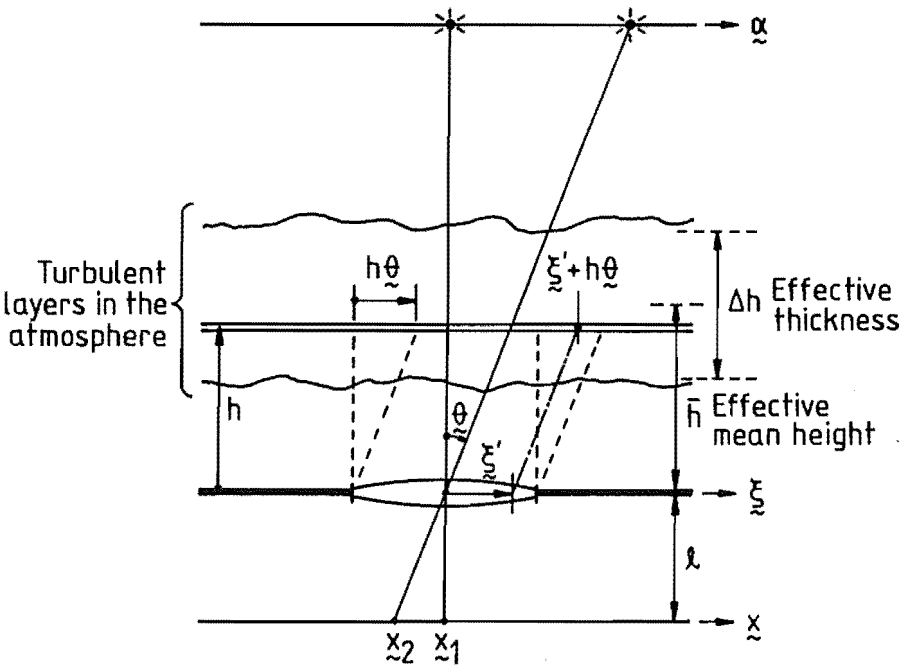


Figure 2.6: Geometry for equations (2.30) and following, describing the effects of partial isoplanatism on the speckle cross-spectrum from two separated celestial point sources.

$B_{LT}(\underset{\sim}{\xi},\lambda)$ is the long exposure atmospheric transfer function (2.3, 2.8e), $Q_4(\underset{\sim}{\xi},\bar{h}\underset{\sim}{\theta})$ is the four-circle overlap function [cf. Korff 1973, Roddier et al 1982 eq. (8)] defined by

$$Q_4(\underset{\sim}{\xi},\underset{\sim}{\xi}') = \int Q(\underset{\sim}{\xi}'')\,Q^*(\underset{\sim}{\xi}+\underset{\sim}{\xi}'')\,Q^*(\underset{\sim}{\xi}'+\underset{\sim}{\xi}'')\,Q(\underset{\sim}{\xi}+\underset{\sim}{\xi}'+\underset{\sim}{\xi}'')\,.d\underset{\sim}{\xi}'' \tag{2.32}$$

for the general case of a complex pupil function Q (see Figure 2.7), and $\sigma(\underset{\sim}{\theta})$ is the correlation area of the perturbed wavefronts ψ at the telescope aperture given by

$$\sigma(\underset{\sim}{\theta}) = \sigma\, e^{-7.65\frac{\Delta h^2}{r_o^2}|\underset{\sim}{\theta}|^2} \tag{2.33}$$

[Roddier et al 1982 eq. (42)] where $\sigma = \sigma(\underset{\sim}{0})$ is given by (2.21). $\bar{h}$ and Δh are the effective mean height and thickness, respectively, of the atmospheric turbulence (Figure 2.6) and are defined in terms of moments with respect to height of the $C_N^2(h)$ profile (ϕ2.2) [Roddier et al 1982 eqs.(37,41)]. If $\underset{\sim}{\theta} = \underset{\sim}{0}$ or $\bar{h} = \Delta h = 0$, i.e. the

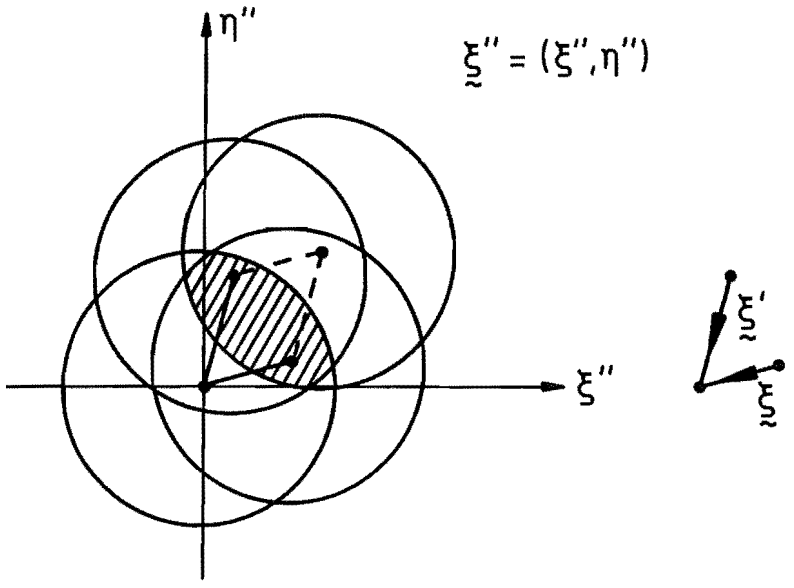


Figure 2.7: The four-circle overlap function $Q_4(\underset{\sim}{\xi},\underset{\sim}{\xi}')$ (2.32) for a circular pupil $Q(\underset{\sim}{\xi})$. If the telescope is unaberrated and unapodised, i.e. $Q(\underset{\sim}{\xi}) = 1$ inside the pupil and 0 outside it, $Q_4(\underset{\sim}{\xi},\underset{\sim}{\xi}')$ is equal to the shaded area.

conditions are isoplanatic, then (2.30) reduces directly to (2.20).

The first term on the right hand side of (2.30) describes the broad background observed in the cross-correlation of two speckle point spread functions. The second term describes the sharp peak that occurs in the cross-correlation under isoplanatic or partially isoplanatic conditions. The width of the peak is inversely related to the width in $\underset{\sim}{\xi}$-space of the product $\sigma(\underset{\sim}{\theta})Q_4(\underset{\sim}{\xi},\bar{h}\underset{\sim}{\theta})$. The $\underset{\sim}{\theta}$-dependence of the shape of $Q_4(\underset{\sim}{\xi},\bar{h}\underset{\sim}{\theta})$ in $\underset{\sim}{\xi}$-space gives rise to different isoplanatic patch sizes for different frequencies and orientations of spatial frequency components in the object [Korff et al 1975 ϕII]. The height of the cross-correlation peak above the background, relative to its height under perfectly isoplanatic conditions, is given by the ratio [cf. Roddier et al 1982 eq.(33)]

$$R_h(\underset{\sim}{\theta}) \approx \frac{\sigma(\underset{\sim}{\theta}) \int Q_4(\underset{\sim}{\xi},\bar{h}\underset{\sim}{\theta}).d\underset{\sim}{\xi}}{\sigma(\underset{\sim}{0}) \int Q_4(\underset{\sim}{\xi},\underset{\sim}{0}).d\underset{\sim}{\xi}} = \frac{\sigma(\underset{\sim}{\theta})}{\sigma(\underset{\sim}{0})} \left(\frac{|Q_2(\bar{h}\underset{\sim}{\theta})|}{|Q_2(\underset{\sim}{0})|}\right)^2 \qquad (2.34)$$

where

$$Q_2(\underset{\sim}{\xi}) = \mathbf{A}[Q(\underset{\sim}{\xi})] = \int Q^*(\underset{\sim}{\xi}')Q(\underset{\sim}{\xi}+\underset{\sim}{\xi}').d\underset{\sim}{\xi}' \qquad (2.35)$$

is the autocorrelation or two-circle overlap function of the pupil function $Q(\underset{\sim}{\xi})$ (Figure 1.8b) and is also the telescope's incoherent transfer function (2.8d). For an unapodised telescope, the ratio of the magnitudes of the Q_2 functions on the right hand side of (2.34) represents the ratio of the overlap area of projections of the telescope's pupil in the directions of the two point sources onto a plane at height $\bar{h}$ to the area of the pupil itself (cf. Figures 1.8, 2.6). Introducing the symbol R_a for this ratio of areas, defined by

$$R_a(\underset{\sim}{\theta}) = \frac{|Q_2(\bar{h}\underset{\sim}{\theta})|}{|Q_2(\underset{\sim}{0})|}, \qquad (2.36)$$

allows (2.34) to be rewritten as

$$R_h(\underset{\sim}{\theta}) \approx \frac{\sigma(\underset{\sim}{\theta})}{\sigma(\underset{\sim}{0})} R_a^{\,2}(\underset{\sim}{\theta}). \qquad (2.37)$$

For thin turbulence ($\Delta h \ll \bar{h}$), $\sigma(\underset{\sim}{\theta})$ is essentially constant in (2.34),

so that (2.37) becomes

$$R_h(\underset{\sim}{\theta}) \approx R_a^{\,2}(\underset{\sim}{\theta}) \tag{2.38}$$

as derived by Lohmann and Weigelt [1979] and Bagnuolo [1984] for a single turbulence layer. $R_h(\underset{\sim}{\theta})$ is a useful definition of the degree of isoplanatism across an angle $\underset{\sim}{\theta}$ in object space because its value ranges from unity in completely isoplanatic conditions (e.g. $\underset{\sim}{\theta} = \underset{\sim}{0}$) to zero in completely nonisoplanatic conditions, and because it describes a tangible and readily measurable feature of the speckle cross-correlation. $R_h(\underset{\sim}{\theta})$ is also readily measurable from the image formed by speckle interferometry (3.14) of a pair of unresolved stars [Roddier et al 1982].

Increasing $|\underset{\sim}{\theta}|$, i.e. reducing the degree of isoplanatism, affects the sharp peak in $\langle H^*(\underset{\sim}{u},\underset{\sim}{x}_1,\nu,t)H(\underset{\sim}{u},\underset{\sim}{x}_2,\nu,t)\rangle_{LT}$ by broadening it and attenuating it relative to its condition for $\underset{\sim}{\theta} = \underset{\sim}{0}$. The size of the isoplanatic patch, which is here taken to be the maximum magnitude θ_o of angular separation for which the sharp peak in the speckle cross-correlation is tangible, depends on the effective width with respect to $\underset{\sim}{\theta}$ of the $\sigma(\underset{\sim}{\theta})\, Q_4(\underset{\sim}{\xi},\bar{h}\underset{\sim}{\theta})$ term in the right hand side of (2.30). The width of $Q_4(\underset{\sim}{\xi},\bar{h}\underset{\sim}{\theta})$ can be described by the value

$$\theta_{oT} = D/\bar{h} \tag{2.39}$$

of $|\underset{\sim}{\theta}|$ above which $Q_4(\underset{\sim}{\xi},\bar{h}\underset{\sim}{\theta})$ is zero for all $\underset{\sim}{\xi}$. The effective width of $\sigma(\underset{\sim}{\theta})$ can be described by the angle $|\underset{\sim}{\theta}| = \theta_{oA}$ at which $\sigma(\underset{\sim}{\theta})$ drops to a fraction $1/e$ of $\sigma(\underset{\sim}{\theta})$ [Roddier et al 1982], which from (2.33) is given by

$$\theta_{oA} = 0.36\, r_o/\Delta h \ . \tag{2.40}$$

The <u>isoplanatic angle</u> θ_o is primarily determined by the smaller of θ_{oT} or θ_{oA}, i.e.

$$\theta_o = \min[\theta_{oT},\ \theta_{oA}] \ , \tag{2.41}$$

with the proviso that the $\sigma(\underset{\sim}{\theta})\, Q_4(\underset{\sim}{\xi},\bar{h}\underset{\sim}{\theta})$ term of (2.30) is exactly zero for $|\underset{\sim}{\theta}| > \theta_{oT}$ but may have significant value for $|\underset{\sim}{\theta}| > \theta_{oA}$.

If $\Delta h/\bar{h} < 0.36 r_o/D$, i.e. the atmospheric turbulence is concentrated in a high, thin layer, then $\theta_{oT} < \theta_{oA}$ and the isoplanatic angle is dependent primarily on the diameter of the telescope aperture, not on the seeing. This is the result obtained by Lohmann and Weigelt [1979] and Bagnuolo [1984] using single phase screen models of the atmosphere. If $\Delta h/\bar{h} > 0.36 r_o/D$, i.e. the turbulence is thick and extends down close to the telescope (which is realistic in practice), then $\theta_{oT} > \theta_{oA}$ and the isoplanatic angle is limited primarily by the atmosphere, independent of telescope size. Putting reasonable values of D = 2m, $\bar{h}$ = 10km into (2.39) [cf. Weigelt 1975 φ5] produces the result $\theta_{oT} \approx 40$ arc seconds, which is larger than the measured values of θ_o quoted in the next paragraph. Thus, in practice the isoplanatic angle of large telescopes should usually be limited by θ_{oA}.

Published observations of nonisoplanatism in short exposure imaging [Title et al 1975, Nisenson and Stachnik 1978, Schneiderman and Karo 1978a, Pollaine et al 1979, Lohmann and Weigelt 1979, Weigelt 1979, Ebersberger and Weigelt 1985] show a wide spread in estimates of θ_o, with Schneiderman and Karo [1978a] reporting no detectable isoplanatism for $|\underline{\theta}| = 3.5"$ and Weigelt [1979] reporting partial isoplanatism for $|\underline{\theta}| = 22"$. This spread may come from differences in atmospheric conditions and from differences in the interpretation of nonisoplanatic effects in the various experiments. Roddier et al [1982] show a graph of θ_{oA} against time, estimated according to (2.40) from repeated measurements of the $C_N^2(h)$ profile, which shows strong fluctuations between about 2" and 5" with a time scale of the order of five to ten minutes. A typical upper limit on θ_o for a detectable peak in the speckle cross-correlation might be taken to be about 10".

2.7 A SIMPLE MODEL OF THE ATMOSPHERE

A feature of speckle processing techniques (Chapter 3), verified by practice in the optical laboratory and in computer simulations [Labeyrie 1976, Bates 1982b], is that in general they are insensitive to the details of the statistics of the distortions to which the incident radiation is subjected. The following simple model is therefore useful in discussing and simulating speckle processing

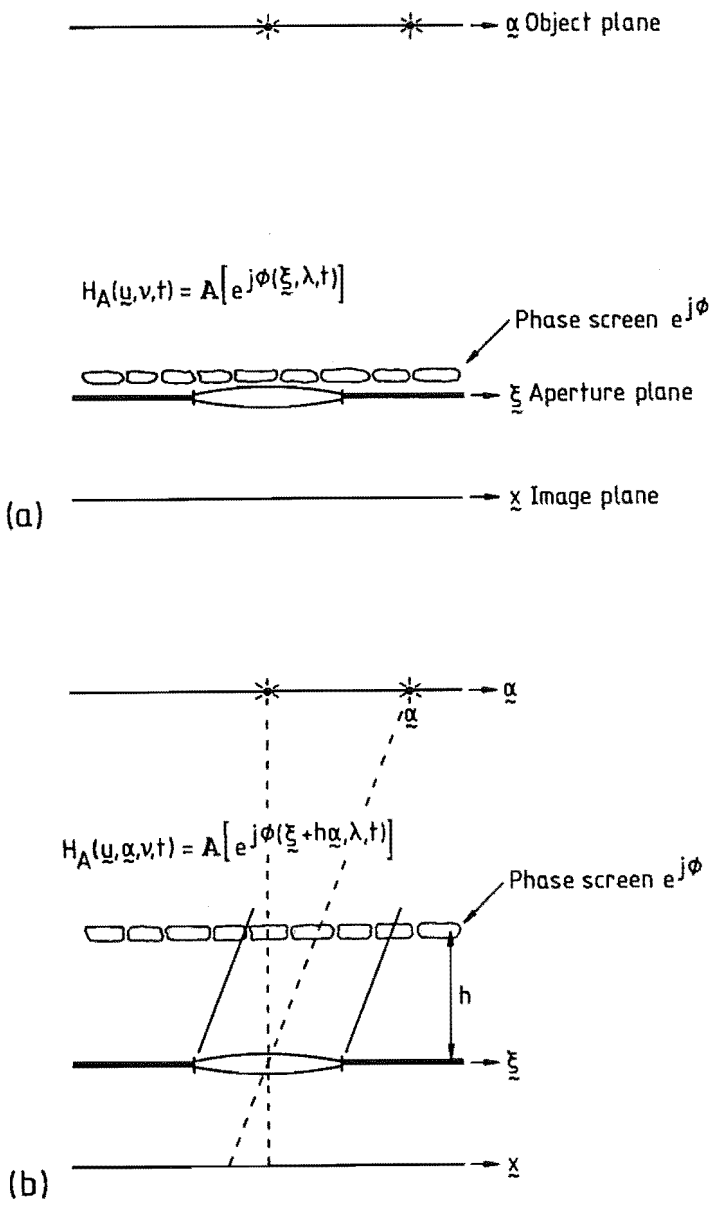


Figure 2.8: A simple model of the atmosphere. The atmosphere is modelled as a thin, transparent "phase screen" which introduces phase perturbations to radiation arriving at the telescope aperture. The phase screen may be pictured as a random collection of "seeing cells" of random refractive index and average diameter r_o [cf. Bates 1982b ϕ4.1].

(a) Placing the phase screen at the aperture plane produces isoplanatic imagery.

(b) Nonisoplanatic imagery may be modelled by raising the phase screen to a height h above the aperture plane.

methods (cf. Chapter 4).

The atmosphere in this model (see Figure 2.8) is represented by a thin, transparent phase screen which randomly perturbs the phase of radiation passing through it. The phase screen is assumed to lie deep within the near field of the telescope aperture (cf. Table 1.2), so that a plane wave of unit amplitude propagating downward through the phase screen arrives at the telescope aperture with effectively unit magnitude and distorted phase. The phase distortion $\phi(\underset{\sim}{\xi},\lambda,t)$ describes the phase perturbation at wavelength λ and time t across that part of the phase screen onto which the telescope aperture is projected by the direction of incidence of the plane wave.

If the phase screen is placed immediately above the telescope aperture (Figure 2.8a), the transfer function for a point object is independent of the position of the object on the celestial sphere (i.e. the imaging system is isoplanatic) and is given by

$$H_A(\underset{\sim}{u},\nu,t) = \mathbf{A}\left[e^{j\phi(\underset{\sim}{\xi},\lambda,t)}\right] . \qquad (2.42)$$

Nonisoplanatic effects are modelled by placing the phase screen at some height h above the telescope aperture (Figure 2.8b), which might perhaps be identified with the peak in the $C_N^2(h)$ profile in the tropopause (cf. Figure 2.2). The atmospheric transfer function for a point object then depends on the object's angular position $\underset{\sim}{\alpha}$ on the celestial sphere and is given by

$$H_A(\underset{\sim}{u},\underset{\sim}{\alpha},\nu,t) = \mathbf{A}\left[e^{j\phi(\underset{\sim}{\xi}+h\underset{\sim}{\alpha},\lambda,t)}\right] . \qquad (2.43)$$

The single phase screen model has been used by Lohmann and Weigelt [1979] and Bagnuolo [1984] to describe the effects of nonisoplanatism on speckle processing techniques. Schneiderman and Karo [1978a] note that the speckle cross-spectrum derived by Korff et al [1975] can be modelled closely using a single phase screen.

2.8 TECHNIQUES FOR COMBATTING THE SEEING

High resolution optical astronomy advanced considerably with the invention of speckle processing techniques (see Chapter 3) which extract high spatial frequency information from ensembles of speckle images (§2.5). However, other techniques for obtaining spatial resolution through the atmosphere beyond the seeing limit of conventional long exposure imagery (§2.4) have been developed which do not operate on ensembles of speckle images and hence are essentially different from speckle processing. Some of these methods are briefly described below. Published reviews include Roddier [1981 §§9.1-9.3] and Bates [1982b §11].

The spatial resolution of long exposure imaging with a circular aperture can be optimised by choosing the aperture diameter appropriately for the seeing and by auto-guiding the telescope on the centroid of a stellar image to remove the image wander component of the seeing disc [Fried 1966, Young 1974]. Image wander is the random fluctuation in position of a stellar image which occurs slowly enough for the eye to follow or a servomechanism to correct. It comes from the randomly varying mean phase gradient of the "instantaneous" coherent transfer function H_c (§2.4) across the telescope pupil. Image wander is pronounced in small telescopes with $D \leq r_o$, becoming less so as the aperture diameter is increased above r_o [Linfoot and Whitcomb 1972, cf. Ohtsubo 1985 §4]. When $D >> r_o$, as is almost always the case with the largest optical telescopes, little image wander occurs and auto-guiding gives negligible improvement to resolution.

Fried [1978] considers the possibility of an individual "lucky" short exposure image being diffraction-limited due to the random atmospheric phase perturbations across the telescope aperture being negligible for the duration of the exposure. He shows theoretically that, for a circular aperture of diameter $D >> r_o$, the probability of this occurring is proportional to

$$e^{-0.16(D/r_o)^2}$$

which decreases rapidly as D increases. Englander et al [1983] confirm this result experimentally. Englander et al note that

treating the best 1% of short exposures as lucky gives a resolution improvement of about six times over the seeing limit, and that nonisoplanatic effects tend to distort the shape of a lucky image of an extended object.

The techniques of image restoration [Bates and McDonnell 198_] may be used to improve spatial resolution by measuring or estimating the long exposure transfer function and then deconvolving it from the long exposure image [Roddier 1981 §9.1]. Because the long exposure transfer function decays rapidly towards zero (see (2.11) and Figure 2.3) the resolution improvement possible in practice, in the presence of images contaminated by noise and recording nonlinearities, is small, subjectively about a factor of two [Worden 1977; cf. McGlamery 1967, Heasley 1984]. McGlamery [1967] considers the use of a reference object to measure the short exposure transfer function and deconvolve it from individual speckle images.

Adaptive optics [Hardy 1978, Roddier 1981 §9.2, Bates 1982b §11.1] seeks to compensate for the optical effects of atmospheric turbulence in real-time by using servomechanisms to introduce time-varying compensating phase delays across the pupil of the telescope, e.g. by deforming the telescope's primary mirror. This technique is in principle very powerful and has been successfully used to sharpen images of bright stars [Pollaine et al 1979]. Its practical application is likely to be restricted to bright objects, or objects within the same isoplanatic patch [Fried 1982] as a bright reference object, since its potential performance is worse than that of postdetection processing (e.g. speckle processing) in the presence of photon noise from dim objects [Roddier and Roddier 1978; cf. Bates 1982b §11.1].

The Michelson stellar interferometer [Michelson and Pease 1921, Hanbury Brown 1974] is a historically important instrument [cf. De Vorkin 1975] which provided the first accurate measurements of the apparent angular diameters of stars other than the Sun. It essentially consists of two separated small telescopes [cf. Labeyrie 1975] which form images in a common image plane. Optical filters are used to make the images quasi-monochromatic [cf. Cutter and Lohmann 1974]. The images from the two telescopes interfere with each other. Under

appropriate conditions interference fringes appear in the composite image. The contrast of the fringes allows the magnitude $|\gamma_s(\underline{\xi})|$ of the mutual coherence (ϕ1.4.2) between the wavefronts entering the telescopes' apertures to be determined [cf. Hanbury Brown 1974 ϕ4.1.1, Born and Wolf 1975 ϕ10.4.3], where $\underline{\xi}$ is the vector separation between the two apertures in the aperture plane (i.e. $\underline{\xi}$ is the interferometer baseline). By the van Cittert-Zernike theorem [Born and Wolf 1975 ϕ10.4.2, Hanbury Brown 1974 ϕ3.3], the complex quantity $\gamma_s(\underline{\xi})$ is proportional to the visibility $F(\underline{u})$ of the celestial object in the absence of atmospheric turbulence (cf. ϕ1.6.4). The atmosphere randomly perturbs the phase of $\gamma_s(\underline{\xi})$ but, under the near-field approximation (ϕ2.2), does not affect its magnitude. Varying the baseline $\underline{\xi}$ allows visibility space to be explored and $|F(\underline{u})|$ to be evaluated. From $|F(\underline{u})|$ the spatial extent of the celestial object, e.g. the diameter of a stellar disc [Hanbury Brown 1974 ϕϕ2.2,10], can be straightforwardly determined (cf. ϕ3.4.2).

The great practical difficulty in using the Michelson stellar interferometer is its stringent requirement on optical path length equality between the two arms of the interferometer for $|\gamma_s(\underline{\xi})|$ to be determinable from the interference fringes. The path length difference must be considerably less than the coherence length of the light forming the images, which for (say) 50nm bandwidth of visible light is about 6μm (cf. 1.33 - 1.36). This requires extreme precision in the fabrication and guiding of the interferometer, the more so as the length of the baseline is increased.

Michelson and Pease in the 1920s estimated the angular diameters of six giant stars [Hanbury Brown 1974 p.25] using baselines up to 6m in length [Hanbury Brown 1974 ϕ2.2, cf. Michelson and Pease 1921], obtaining a potential angular resolution of 0.02 seconds of arc. Subsequent attempts by Pease to obtain better resolution with a 15m baseline were unsuccessful [De Vorkin 1975]. The modern status of Michelson stellar interferometry is reviewed by Tango and Twiss [1980] and briefly by Bates [1982b ϕ11.6]. Labeyrie [1982] describes recent progress with coherent arrays of optical telescopes.

The intensity interferometer of Hanbury Brown and Twiss

[Hanbury Brown 1974, cf. Hanbury Brown and Twiss 1954] electrically correlates the intensity fluctuations of light arriving at two separated apertures. The correlation of the intensity fluctuations is ideally equal to $|\gamma_s(\underset{\sim}{\xi})|^2$ [Hanbury Brown 1974 ϕϕ3.5,4.2.2], yielding the same information, $|F(\underset{\sim}{u})|$, as does the Michelson stellar interferometer. The intensity interferometer is very much less sensitive than the Michelson stellar interferometer to optical path length differences between the interferometer arms, because its effective optical bandwidth is limited by the electrical bandwidth of the correlator to a very small value, e.g. $\Delta\lambda \approx 0.0001$nm in the visible spectrum for 100MHz electrical bandwidth. Hanbury Brown [1974 ϕ2.3] quotes 30cm as an upper limit on path length differences, which is easy to achieve in practice. The small effective optical bandwidth and the multiplying together of the photodetector outputs [Hanbury Brown 1974 ϕ4.2.1] in the correlator give the intensity interferometer an inherently poor signal-to-noise ratio. This restricts its use to comparatively bright celestial objects.

The intensity interferometer at Narrabri in Australia [Hanbury Brown 1974 ϕ8], which had a maximum baseline of 188m, provided measurements of the angular diameters of 32 stars other than the six red giants previously measured by Michelson and Pease [Hanbury Brown et al 1974, Hanbury Brown 1974 pp.134,135,163], achieving an angular resolution of 0.4 milliseconds of arc. Hanbury Brown and his colleagues have since abandoned intensity interferometry in favour of Michelson interferometry for high resolution studies because of the better signal-to-noise ratio of the latter technique [Hanbury Brown 1984].

Pupil plane interferometry [cf. Greenaway and Dainty 1978, Tango 1979] is a comparatively recent development in which two images of the pupil of a telescope are superimposed, displaced with respect to each other by reflection [Dainty and Scaddan 1974], translation [Brown and Scaddan 1979, Ribak and Leibowitz 1985, Ribak et al 1985b], or rotation [Roddier 1976, Breckinridge 1978] in the pupil plane, so that they interfere with each other. This is in contrast to conventional image plane interferometry, such as Michelson stellar interferometry, in which images in the image plane (i.e. the spatial Fourier transform of the pupil plane (ϕ1.6.3)) are superimposed.

Pupil plane interferometry has the advantages over image plane techniques of better signal-to-noise ratio [Greenaway and Dainty 1978], since it is unaffected by pupil redundancy (ϕ1.6.5) [Roddier and Roddier 1978] and can be used with wideband light, and the ability to measure the magnitude of a celestial object's visibility without needing corrections for atmospheric turbulence or telescope aberrations [Roddier and Roddier 1983].

The rotation-shearing pupil plane interferometer described by Roddier [1976] is a compact instrument [Roddier 1982] which is used with a conventional telescope. The "image" it produces provides a two-dimensional map of the mutual coherence $\gamma_s(\underline{\xi})$ of the radiation at the telescope aperture [Roddier 1976]. This information is equivalent to that provided by many Michelson interferometers working simultaneously with many different baselines $\underline{\xi}$. Viewing a stellar object with the object displaced from the optical axis of the telescope produces a series of fine fringes in the "image" [Roddier and Roddier 1978]. The contrast of these fringes, which is not affected by fixed aberrations in the optics, yields $|\gamma_s(\underline{\xi})|$ straightforwardly. For fixed aberrations the position of the fringes together with observations of a reference object allows the phase of $\gamma_s(\underline{\xi})$ to be determined and hence an image of the object to be formed [Roddier and Roddier 1978, Roddier and Breckinridge 1984]. When viewing through the atmosphere, the randomly varying aberrations due to atmospheric turbulence can be effectively frozen by taking a short exposure [Roddier 1976]. $|\gamma_s(\underline{\xi})|$ is then obtained with no need to correct for effects of the atmosphere or telescope aberrations [Roddier and Roddier 1983]. Reconstruction of an image of the celestial object may be attempted [Roddier and Roddier 1985] by Fourier phase retrieval (ϕ3.4.2).

2.9 SUMMARY

The random phase perturbations induced by the atmosphere in optical wavefronts arriving at the telescope are a serious hinderance to high resolution astronomical imaging. An important measure of the quality of the seeing is the Fried parameter r_o, which describes the correlation length of the atmospheric phase perturbations in the

aperture plane of the telescope. High spatial frequencies are lost in conventional long exposure images because the long exposure OTF of a large telescope and the atmosphere combined decays rapidly with increasing spatial frequency $|\underline{u}|$ and is essentially zero above the seeing limit $|\underline{u}| = r_o/\lambda\ell$. Short exposure, narrowband speckle images retain high spatial frequency information out to the diffraction limit of the telescope, although this information is distorted by a transfer function which has random phase. Nonisoplanatism reduces the correlation between speckle images of spatially separated objects. Various techniques for obtaining spatial resolution beyond the seeing limit have been developed, many of which yield only the magnitude of the visibility of an object and hence do not directly produce an image of the object. Speckle processing techniques, which operate on ensembles of speckle images, are reviewed in Chapter 3.

CHAPTER 3

ASTRONOMICAL SPECKLE PROCESSING TECHNIQUES

3.1 INTRODUCTION

High spatial resolution in optical astronomy can be obtained with a conventional telescope by suitably processing short exposure speckle images (ϕ2.5). As explained in Chapter 2, speckle images contain information about the celestial object at spatial frequencies out to the diffraction limit of the telescope. Whereas the high spatial frequencies are blurred and lost in the long exposure image (ϕ2.4), they are present in speckle images though attenuated and distorted by the random short exposure transfer function of the atmosphere (ϕ2.5).

In 1970, Labeyrie introduced the first of a class of techniques for obtaining high spatial resolution in optical astronomy. This class is here referred to collectively as speckle processing techniques. These are distinguished from other high resolution techniques (ϕ2.8), such as Michelson stellar interferometry, coherence interferometry, or adaptive optics, by operating on ensembles of speckle images. The underlying assumptions of speckle processing techniques are that speckle images exposed at instants separated by more than the redistribution time of the atmosphere (ϕ2.2) experience effectively independent random atmospheric distortions, and that over a certain field of view the distortion in each speckle image is effectively isoplanatic. By combining many statistically independent speckle images of an object, speckle processing techniques aim to recover information about the object at spatial frequencies out to the diffraction limit of the telescope.

Speckle processing techniques can be divided into those which form an estimate of the true image (ϕ3.2) of the object and those which form something else, e.g. an image resembling the autocorrelation of the object. Labeyrie's original technique, which has come to be known as speckle interferometry, does not estimate

the true image, as is explained in Section 3.3. The name speckle imaging is here applied to speckle processing techniques which do estimate the true image. Papaliolios et al [1985] apply the name to the Knox-Thompson method alone (§3.5.2), but there are many other methods (§§ 3.4 - 3.6) which can successfully estimate the true image and I believe the name is better used in the general sense. The term "speckle interferometry" is used outside astronomy to describe a branch of the science of speckle metrology [Ennos 1984, Erf 1978], but in this thesis only the astronomical meaning is intended.

This chapter reviews speckle processing techniques in optical astronomy. The mathematical notation of Chapters 1 and 2 is trimmed for the sake of brevity and some new terminology is introduced in Section 3.2. Speckle interferometry is explained in Section 3.3, with the aid of one-dimensional computer-generated examples (cf. §4) which also serve to illustrate the discussions of random atmospheric transfer functions in this introduction and in Sections 2.4, 2.5, and 2.9 of the previous chapter. The various methods proposed to compensate for the transfer function of speckle interferometry are examined in Section 3.3.2. Section 3.4 describes speckle imaging methods which estimate the true image from the autocorrelation estimate generated by speckle interferometry. Speckle imaging methods which use speckle images directly are covered in Sections 3.5 and 3.6. Section 3.5 describes methods that operate in the Fourier (or visibility) domain, and Section 3.6 describes methods that operate in the image domain, i.e. the shift-and-add methods. Section 3.7 presents a brief summary.

Many speckle processing techniques have been devised in the sixteen years since Labeyrie [1970] introduced speckle interferometry. This chapter does not describe all of them. Notable omissions are Beckers's differential speckle interferometry [Beckers and Hege 1981, Beckers 1982; cf. Aime et al 1984, Hebden et al 1985, Hege et al 1985, Petrov et al 198_] and Weigelt et al's [1985b] speckle spectroscopy, neither of which have estimation of the true image as their primary objective. Other astronomical imaging proposals which are merely mentioned in this introduction are those of Greenaway [1982] and Huiser [1982b]. These are sometimes quoted in speckle processing contexts, but they are not speckle processing techniques

according to the definition above because they do not operate on ensembles of speckle images.

Published review articles describing astronomical speckle processing include Dainty [1975b], Worden [1977], Bates [1982b], and Dainty [1984] which give comprehensive coverage of techniques (see also Bracewell [1979]), and Labeyrie [1976], McAlister [1977], Labeyrie [1978], and Nisenson et al [1983a] which describe particular techniques.

3.2 NOTATION AND TERMINOLOGY

The notation developed in Chapters 1 and 2 leads to the description

$$s_{SE}(\underset{\sim}{x},\nu,t) = f(\underset{\sim}{x},\nu) \circledcirc h(\underset{\sim}{x},\nu,t) \tag{3.1}$$

(2.13e) for a speckle image, in which the dependence of s_{SE}, f, and h on time t and optical frequency ν appears explicitly. Speckle images are customarily formed with narrowband optical filters to improve the speckle contrast and hence improve the recovery of high spatial frequencies (cf. §2.5) by speckle processing techniques. The object $f(\underset{\sim}{x},\nu)$ is then effectively quasi-monochromatic, and the mean optical frequency $\bar{\nu}$ or wavelength $\bar{\lambda}$ can be taken as understood. Also, the exposure time of the speckle images is usually treated as being so short that the atmosphere remains effectively unchanging or frozen during the exposure. Each speckle image can thus be considered to be exposed at an instant t_m and to be the m^{th} member of an ensemble of M such speckle images obtained while observing a particular object. The m^{th} speckle image of the quasi-monochromatic object $f(\underset{\sim}{x})$ (with $\bar{\nu}$ understood) is denoted $s_m(\underset{\sim}{x})$ and has an associated point spread function $h_m(\underset{\sim}{x})$, these quantities being related to the quantities in (3.1) by

$$\begin{aligned} s_m(\underset{\sim}{x}) &= s_{SE}(\underset{\sim}{x},\bar{\nu},t_m), \\ f(\underset{\sim}{x}) &= f(\underset{\sim}{x},\bar{\nu}), \\ h_m(\underset{\sim}{x}) &= h(\underset{\sim}{x},\bar{\nu},t_m). \end{aligned} \tag{3.2}$$

The ensemble of speckle images is denoted $\{s_m(\underset{\sim}{x})\}$. The individual members of the ensemble are random quantities that are treated in this thesis as being mutually statistically independent, i.e. the speckle images are assumed to be exposed at instants separated by more than the redistribution time of the atmosphere (§2.2). This treatment may be extended to speckle images exposed at instants separated by less than the redistribution time (as is often the case in observational practice) by considering the effective value of M to be reduced according to the degree of statistical dependence between consecutive speckle images [Bates 1982b §4.4]. The terminology associated with the quantities of (3.2) and their spatial Fourier transforms is as follows:

object	$f(\underset{\sim}{x})$	$\leftrightarrow$	$F(\underset{\sim}{u})$	object visibility,
speckle psf	$h_m(\underset{\sim}{x})$	$\leftrightarrow$	$H_m(\underset{\sim}{u})$	speckle transfer function,
speckle image	$s_m(\underset{\sim}{x})$	$\leftrightarrow$	$S_m(\underset{\sim}{u})$	speckle visibility.

The equations (2.13) describing short exposure imaging are restated in the notation of (3.2) as

$$\underset{\sim}{\xi} = \bar{\lambda}\ell\underset{\sim}{u}, \tag{3.3a}$$

$$H_{cm}(\underset{\sim}{u}) = Q(\bar{\lambda}\ell\underset{\sim}{u})\,\psi_m(\bar{\lambda}\ell\underset{\sim}{u}), \tag{3.3b}$$

$$H_m(\underset{\sim}{u}) = \mathbf{A}[H_{cm}(\underset{\sim}{u})], \tag{3.3c}$$

$$h_m(\underset{\sim}{x}) \leftrightarrow H_m(\underset{\sim}{u}), \tag{3.3d}$$

$$s_m(\underset{\sim}{x}) = f(\underset{\sim}{x}) \odot h_m(\underset{\sim}{x}) \tag{3.3e}$$

$$S_m(\underset{\sim}{u}) = F(\underset{\sim}{u})\,H_m(\underset{\sim}{u}). \tag{3.3f}$$

There are also the atmospheric transfer function

$$H_A(\underset{\sim}{u}) = H_A(\underset{\sim}{u},\bar{\nu}) \tag{3.4}$$

(cf. 2.8e), the diffraction-limited transfer function

$$H_D(\underset{\sim}{u}) = \mathbf{A}[P(\bar{\lambda}\ell\underset{\sim}{u})] \tag{3.5}$$

(cf. 2.23) with its associated point spread function $h_D(\underset{\sim}{x}) \leftrightarrow H_D(\underset{\sim}{u})$,

and the diffraction-limited image or true image

$$s_D(\underset{\sim}{x}) = f(\underset{\sim}{x}) \odot h_D(\underset{\sim}{x}). \tag{3.6}$$

Ensemble averages over the M quantities in an ensemble are denoted by angle brackets with a subscript M,

$$\langle \cdot \rangle_M = \frac{1}{M} \sum_{m=1}^{M} \cdot_m . \tag{3.7}$$

Unsubscripted angle brackets represent the limit of the average as M tends to infinity,

$$\langle \cdot \rangle = \lim_{M\to\infty} \frac{1}{M} \sum_{m=1}^{M} \cdot_m , \tag{3.8}$$

which is equivalent to the long time average $\langle \cdot \rangle_{LT}$ of Chapter 2. Using (3.8) one has, for example, the long exposure transfer function

$$H_{LE}(\underset{\sim}{u}) = \langle H_m(\underset{\sim}{u}) \rangle \tag{3.9}$$

(cf. 2.8f) and the long exposure image

$$s_{LE}(\underset{\sim}{x}) = \langle s_m(\underset{\sim}{x}) \rangle = f(\underset{\sim}{x}) \odot \langle h_m(\underset{\sim}{x}) \rangle \tag{3.10}$$

(cf. 2.8h).

It is helpful when discussing the results of speckle interferometry and of Fourier phase retrieval to introduce the concept of the form of an object or image, which describes its shape or appearance irrespective of translations in position, changes of brightness, or mirror inversions in image space. Thus, for a given function $f(\underset{\sim}{x})$,

$$f(\underset{\sim}{x}) \sim a f(\underset{\sim}{x}) \sim f(\underset{\sim}{x}-\underset{\sim}{b}) \sim f(-\underset{\sim}{x}) \tag{3.11}$$

where a is an arbitrary scalar constant, $\underset{\sim}{b}$ is an arbitrary vector constant, and "~" is read as "has the same image-form as".

3.3 SPECKLE INTERFEROMETRY

3.3.1 LABEYRIE

Labeyrie [1970] proposes that high spatial resolution can be obtained from speckle images by averaging the power spectra, or squared visibility magnitudes, of the speckle images. This averaging yields

$$S_{LA}(\underset{\sim}{u}) = \langle |S_m(\underset{\sim}{u})|^2 \rangle_M = \langle |F(\underset{\sim}{u}) H_m(\underset{\sim}{u})|^2 \rangle_M$$
$$= |F(\underset{\sim}{u})|^2 \langle |H_m(\underset{\sim}{u})|^2 \rangle_M , \qquad (3.12)$$

which is the product of the power spectrum of the object and the averaged power spectrum, or Wiener spectrum, of the speckle transfer function (3.3c). The expected value of the latter quantity is here called the Labeyrie transfer function and for a large telescope ($D >> r_o$) is given by (cf. 2.20)

$$H_{LA}(\underset{\sim}{u}) = \langle |H_m(\underset{\sim}{u})|^2 \rangle \approx |H_{LE}(\underset{\sim}{u})|^2 + \frac{A}{N_{sp}} H_D(\underset{\sim}{u})$$
$$= A^2 \left(|\hat{H}_{LE}(\underset{\sim}{u})|^2 + \frac{1}{N_{sp}} \hat{H}_D(\underset{\sim}{u}) \right) , \qquad (3.13)$$

where $A = H_{LE}(\underset{\sim}{0}) = H_D(\underset{\sim}{0}) = H_m(\underset{\sim}{0})$ is the area of the telescope aperture and N_{sp} is the average number of speckles per speckle image (2.14). (3.13) contains a component proportional to the diffraction-limited transfer function of the telescope and is thus significantly greater than zero out to the diffraction limit (see Figure 3.1). This allows diffraction-limited information about the object to be obtained from $\langle |S_m(\underset{\sim}{u})|^2 \rangle_M$, even though it cannot be obtained from the long exposure image $\langle s_m(\underset{\sim}{x}) \rangle_M \leftrightarrow \langle S_m(\underset{\sim}{u}) \rangle_M$ (see Figure 3.2, which illustrates Labeyrie's technique as described throughout this subsection). The phases of the $S_m(\underset{\sim}{u})$ are discarded in the averaging process, which allows the magnitude of $F(\underset{\sim}{u})$ to be determined to the diffraction limit at the expense of the phase of $F(\underset{\sim}{u})$.

Speckle interferometry is described in the image domain by taking the inverse Fourier transform of both sides of (3.12). Appealing to the autocorrelation theorem [Bracewell 1978 p.115], this is equal to

$$s_{LA}(\underset{\sim}{x}) = \langle \mathbf{A}[s_m(\underset{\sim}{x})] \rangle_M = \mathbf{A}[f(\underset{\sim}{x})] \odot \langle \mathbf{A}[h_m(\underset{\sim}{x})] \rangle_M \,, \tag{3.14}$$

showing that Labeyrie's procedure is equivalent to averaging the autocorrelations of the speckle images which yields the autocorrelation of the object convolved with a point spread function.

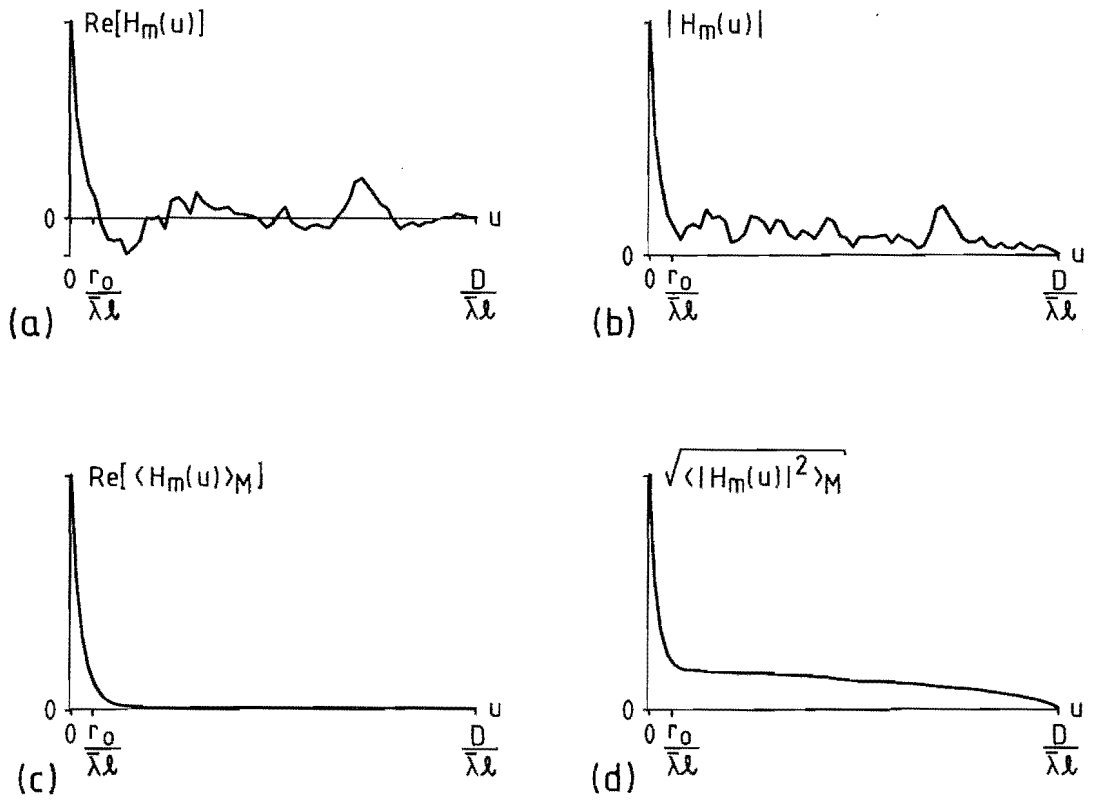


Figure 3.1: One-dimensional computer simulations illustrating the transfer functions in speckle interferometry.

(a) A typical speckle transfer function (the real part is shown).

(b) The magnitude of the speckle transfer function whose real part is shown in (a).

(c) The average of M = 1000 statistically independent realisations of (a). This tends to the long exposure transfer function as $M \to \infty$.

(d) The RMS average of M = 1000 statistically independent realisations of (b). This tends to the square root of the Labeyrie transfer function as $M \to \infty$.

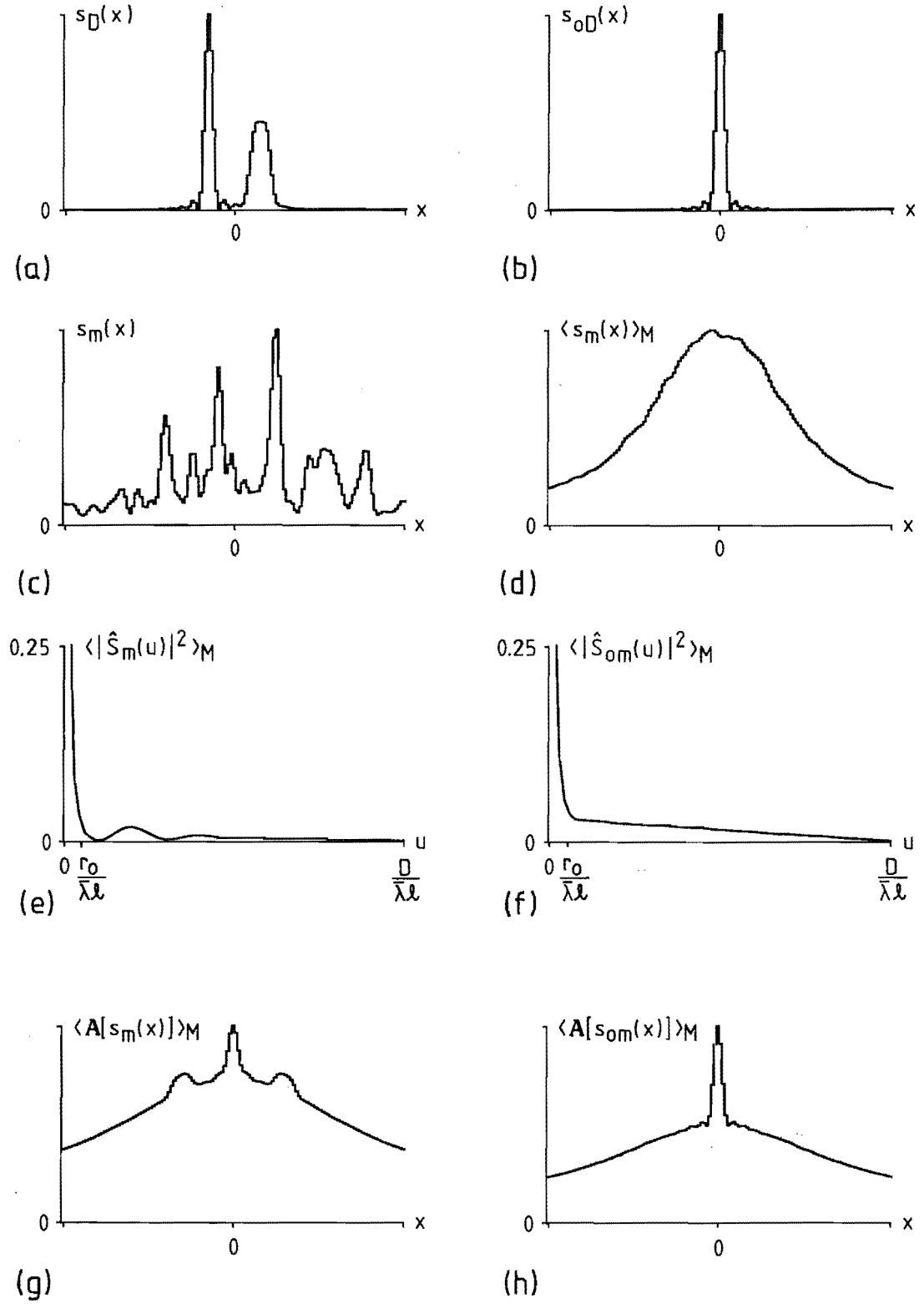


Figure 3.2: One-dimensional computer simulations demonstrating Labeyrie's [1970] speckle interferometry.

(a) True (diffraction-limited) image of an object $f(\underset{\sim}{x})$.

(b) True image of a point object $f_o(\underset{\sim}{x})$.

(c) Typical speckle image of $f(\underset{\sim}{x})$.

(d) Long exposure image of $f(\underset{\sim}{x})$.

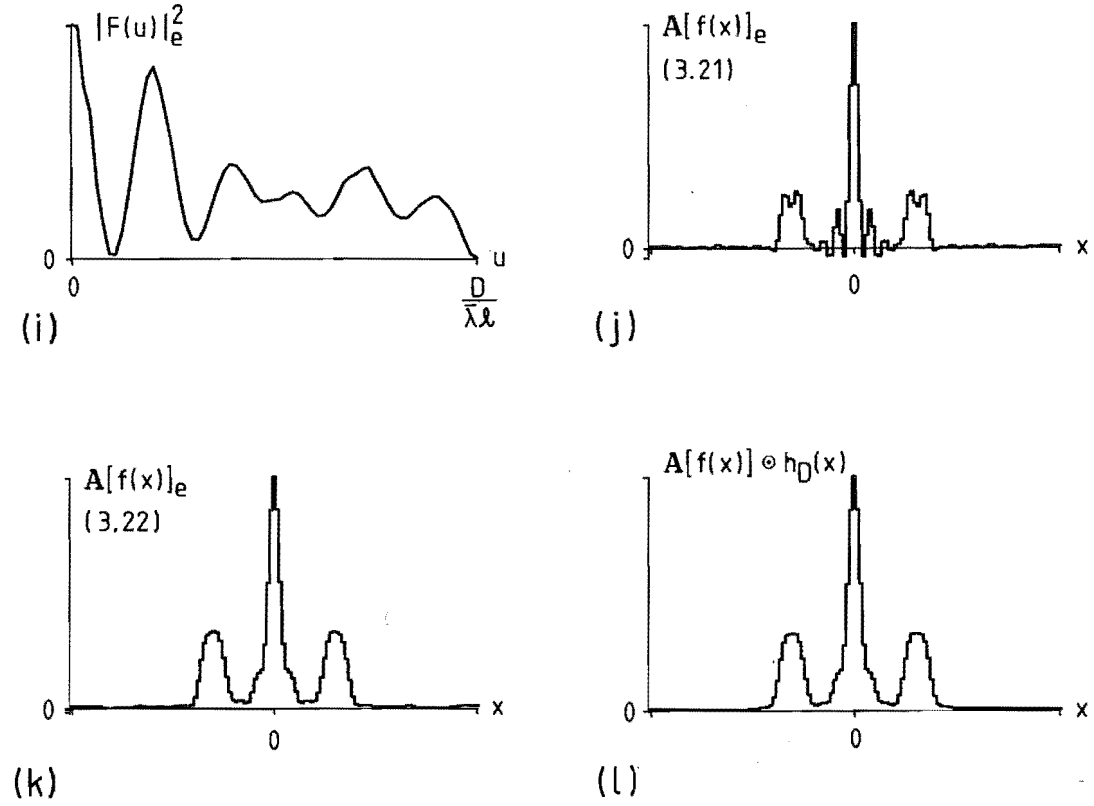


Figure 3.2 (continued):

(e) Average power spectrum of M = 1000 statistically independent realisations of (c), normalised to unity at the origin.

(f) Average power spectrum of M = 1000 statistically independent speckle images of $f_o(\underset{\sim}{x})$ viewed through seeing having identical statistical properties to the seeing of (e), normalised in the same way as (e).

(g) The inverse Fourier transform of (e).

(h) The inverse Fourier transform of (f).

(i) The result of Wiener filtering (f) from (e). This is the speckle interferometry estimate of the power spectrum of the object.

(j) The inverse Fourier transform of (i). This is the speckle interferometry estimate (3.21) of the autocorrelation of the object.

(k) The inverse Fourier transform of the product of (i) and the diffraction-limited transfer function. This is an improved speckle interferometry estimate (3.22) of the autocorrelation of the object.

(l) The true image of the autocorrelation of the object, for comparison with (j) and (k).

The expected value of the point spread function is given by

$$h_{LA}(\underset{\sim}{x}) = \langle \mathbf{A}[h_m(\underset{\sim}{x})] \rangle \tag{3.15}$$

and is here called the Labeyrie point spread function. The diffraction-limited component in $H_{LA}(\underset{\sim}{u})$ corresponds to a spatially narrow component in $h_{LA}(\underset{\sim}{x})$ of the same form as the telescope's diffraction-limited psf $h_D(\underset{\sim}{x})$.

The Labeyrie transfer function is deterministic for given seeing conditions and therefore can be measured or predicted. The resulting estimate of $H_{LA}(\underset{\sim}{u})$ can be used to filter $\langle |H_m(\underset{\sim}{u})|^2 \rangle_M$ from $S_{LA}(\underset{\sim}{u})$ leaving an estimate of $|F(\underset{\sim}{u})|^2$. The standard way of estimating $H_{LA}(\underset{\sim}{u})$ is to perform speckle interferometry on speckle images of an unresolved reference object

$$f_o(\underset{\sim}{x}) = a\,\delta(\underset{\sim}{x}) \tag{3.16}$$

(where a is a constant) observed through seeing statistically similar to that through which $f(\underset{\sim}{x})$ was observed, thereby forming

$$\begin{aligned} S_{oLA}(\underset{\sim}{u}) = \langle |S_{om}(\underset{\sim}{u})|^2 \rangle_{Mo} &= |F_o(\underset{\sim}{u})|^2 \langle |H_{om}(\underset{\sim}{u})|^2 \rangle_{Mo} \\ &= a \langle |H_{om}(\underset{\sim}{u})|^2 \rangle_{Mo} . \end{aligned} \tag{3.17}$$

The subscripts o in (3.17) denote the observation of an unresolved reference object. A filter for $\langle |H_m(\underset{\sim}{u})|^2 \rangle_M$ is formed from $S_{oLA}(\underset{\sim}{u})$, for example the inverse filter

$$S_{oLA}(\underset{\sim}{u})^{-1} = \frac{1}{\langle |S_{om}(\underset{\sim}{u})|^2 \rangle_{Mo}} \approx \frac{1}{a \langle |H_{om}(\underset{\sim}{u})|^2 \rangle_{Mo}} \tag{3.18}$$

or the Wiener filter

$$s_{oLA}(\underset{\sim}{u})^{-1} = \frac{\langle |S_{om}(\underset{\sim}{u})|^2 \rangle^*_{Mo}}{\left(\langle |S_{om}(\underset{\sim}{u})|^2 \rangle_{Mo}\right)^2 + \Phi} \tag{3.19}$$

[Castleman 1979 ϕ11, Bates and McDonnell 198_ ϕϕ16,36]. The filter is applied to $S_{LA}(\underset{\sim}{u})$ to form the speckle interferometry estimate $|F(\underset{\sim}{u})|^2_e$ of the object's power spectrum,

$$|F(\underset{\sim}{u})|_e^2 = S_{LA}(\underset{\sim}{u})\, S_{oLA}(\underset{\sim}{u})^{-1} = \frac{\langle |S_m(\underset{\sim}{u})|^2 \rangle_M}{\langle |S_{om}(\underset{\sim}{u})|^2 \rangle_{Mo}} \quad \text{using (3.18)}$$

$$\approx |F(\underset{\sim}{u})|^2 \frac{\langle |H_m(\underset{\sim}{u})|^2 \rangle_M}{a \langle |H_{om}(\underset{\sim}{u})|^2 \rangle_{Mo}}$$

$$\approx \frac{1}{a} |F(\underset{\sim}{u})|^2 . \qquad (3.20)$$

The approximation in the last line of (3.20) holds below the diffraction limit provided $\langle |H_{om}(\underset{\sim}{u})|^2 \rangle_{Mo}$ is similar to $\langle |H_m(\underset{\sim}{u})|^2 \rangle_M$. Above the diffraction limit, $\langle |H_m(\underset{\sim}{u})|^2 \rangle_M$ is zero and so these spatial frequencies of $|F(\underset{\sim}{u})|^2$ are lost.

The inverse Fourier transform of $|F(\underset{\sim}{u})|_e^2$ gives the speckle interferometry estimate

$$\mathbf{A}[f(\underset{\sim}{x})]_e = \mathbf{F}^{-1}[|F(\underset{\sim}{u})|_e^2] \qquad (3.21)$$

of the autocorrelation of the object, which (provided $\langle |H_{om}(\underset{\sim}{u})|^2 \rangle_{Mo}$ is similar to $\langle |H_m(\underset{\sim}{u})|^2 \rangle_M$) resembles the autocorrelation $\mathbf{A}[f(\underset{\sim}{x})]$ of $f(\underset{\sim}{x})$ convolved with a point spread function like a coherent, not an incoherent, psf (cf. Figure 1.10, ϕϕ 1.6.3, 1.6.4). Inverse Fourier transforming the product of $|F(\underset{\sim}{u})|_e^2$ and the diffraction-limited incoherent transfer function of the telescope [cf. Baba et al 1984] produces an estimate

$$\mathbf{A}[f(\underset{\sim}{x})]_e = \mathbf{F}^{-1}[|F(\underset{\sim}{u})|_e^2 H_D(\underset{\sim}{u})] \qquad (3.22)$$

which ideally has the form of the true image of the autocorrelation of $f(\underset{\sim}{x})$.

Beginning with the initial observations of Gezari et al [1972], speckle interferometry has been part of observational astronomical practice for nearly fifteen years. A glance at the subject index of the Astronomy and Astrophysics Abstracts, e.g. Volume 35/36 page 805 [Bohme et al 1984], shows the wide variety of celestial objects on which speckle interferometry has been used (see also Bates [1982b ϕ9], Dainty [1984 ϕ7.6]). Speckle interferometry readily yields accurate estimates of the angular size and separation of sufficiently

bright objects such as stellar discs [e.g. Balega et al 1982], binary stars [e.g. McAlister 1977, McAlister and Hartkopf 1984], planets and their satellites [e.g. Hege et al 1982a, Baier and Weigelt 1984, cf. McCarthy et al 1985], and asteroids [e.g. Baier and Weigelt 1983, Drummond et al 1985a, 1985b]. Speckle interferometry in the infrared is described by (for example) Sibille et al [1979] and Dyck and Howell [1985]. The need to properly compensate for the Labeyrie transfer function makes accurate photometric information, such as the limb darkening of stellar discs or the relative brightnesses of the components of binary stars, more difficult to obtain than positional information. The difficulty of accurate compensation for the Labeyrie transfer function [cf. Christou 1985a] has slowed the acceptance of speckle interferometry by the astronomical community [Dainty 1984 §7.6].

3.3.2 COMPENSATING FOR THE LABEYRIE TRANSFER FUNCTION

As shown in the previous section, speckle interferometry produces a version of the object's power spectrum (3.12) or autocorrelation (3.14) modified by the Labeyrie transfer function $H_{LA}(\underset{\sim}{u})$ (3.13) which attenuates high spatial frequencies relative to low spatial frequencies. The effect of $H_{LA}(\underset{\sim}{u})$ in image space is shown in Figures 3.2g and h, in which the diffraction-limited autocorrelation of the object appears on top of a broad background or fog. The standard way of estimating $H_{LA}(\underset{\sim}{u})$, so that its effects can be compensated for, is to observe an unresolved object as mentioned in the previous section. For this to be effective the statistical properties of the seeing must be spatially and temporally stationary between the positions and times of observation of the object and the reference object. It is often difficult to achieve this in practice. An extreme case occurs in solar observations [e.g. Ricort and Aime 1979] for which there are no unresolved objects visible at all. This section describes techniques which have been used or proposed to compensate for $\langle|H_m(\underset{\sim}{u})|^2\rangle_M$ in the speckle interferometry observation $S_{LA}(\underset{\sim}{u})$ without recourse to observations of an unresolved reference object.

The Labeyrie transfer function can be predicted from theoretical models of optical propagation through the atmosphere,

given knowledge of the telescope's optical performance and of the Fried parameter r_o (φ2.4) which describes the seeing. r_o can be determined from observations of an arbitrary object, provided the object contains spatial frequencies above the seeing limit, by forming ratios in Fourier space of observations of the object made under different conditions so that the unknown object visibility cancels out of the ratios. Aime et al [1978b] form the ratio

$$\frac{S_{1LA}(\underset{\sim}{u})}{S_{2LA}(\underset{\sim}{u})} = \frac{\langle |H_{1m}(\underset{\sim}{u})|^2 \rangle_{M1}}{\langle |H_{2m}(\underset{\sim}{u})|^2 \rangle_{M2}} \tag{3.23}$$

with the subscripts 1 and 2 indicating observations made through seeing with significantly different values of r_o. von der Lühe [1984] forms the ratio

$$\frac{|S_{LE}(\underset{\sim}{u})|^2}{S_{LA}(\underset{\sim}{u})} = \frac{|\langle H_m(\underset{\sim}{u}) \rangle_M|^2}{\langle |H_m(\underset{\sim}{u})|^2 \rangle_M} , \tag{3.24}$$

which is the ratio of the squared magnitude of the long exposure image to the speckle interferometry image for a single observation. The ratios are solved for r_o using the theory of Fried [1966] and Korff [1973]. Given r_o, $H_{LA}(\underset{\sim}{u})$ is predicted using Korff's [1973] model. Aime et al [1978a] and Ricort and Aime [1979] use this technique in measuring the power spectrum of solar granulations.

Worden et al [1977] (see also Welter and Worden [1978]) propose that the fog in $s_{LA}(\underset{\sim}{x})$ can be removed by subtracting from $s_{LA}(\underset{\sim}{x})$ the averaged cross-correlation $\langle s_m(\underset{\sim}{x}) \otimes s_n(\underset{\sim}{x}) \rangle_M$ of pairs of statistically independent speckle images from the ensemble. This forms an estimate

$$\mathbf{A}[f(\underset{\sim}{x})]_e = \langle s_m(\underset{\sim}{x}) \otimes s_m(\underset{\sim}{x}) \rangle_M - \langle s_m(\underset{\sim}{x}) \otimes s_n(\underset{\sim}{x}) \rangle_M \tag{3.25}$$

which closely resembles the true image of $\mathbf{A}[f(\underset{\sim}{x})]$ provided the angular extent of the object is small compared with that of the seeing disc [Welter and Worden 1978]. Because the speckles in independent speckle images are uncorrelated in position, spatial frequencies above the seeing limit do not survive in the average cross-correlation. $\langle s_m(\underset{\sim}{x}) \otimes s_n(\underset{\sim}{x}) \rangle_M$ is therefore an estimate of the fog [cf. Worden et al

1977 Fig.1]. Worden and Stein [1979] note that the brightness and position of each of the speckle images in the ensemble should be normalised to give autocorrelations and cross-correlations with the same position and brightness in image space. Fante [1979] notes that spatial frequencies below the seeing limit do survive in $\langle s_m(\underset{\sim}{x}) \otimes s_n(\underset{\sim}{x})\rangle_M$ and are therefore strongly attenuated in (3.25), so that the Welter-Worden method is not applicable to objects whose extent is comparable to or larger than that of the seeing disc.

Bruck and Sodin [1980] state that the average cross-correlation of statistically independent speckle images is equal to the autocorrelation of the long exposure image, i.e.

$$\langle s_m(\underset{\sim}{x}) \otimes s_n(\underset{\sim}{x})\rangle = \langle s_m(\underset{\sim}{x})\rangle \otimes \langle s_m(\underset{\sim}{x})\rangle, \tag{3.26}$$

so that in Fourier space (3.25) is equivalent to

$$\begin{aligned} |F(\underset{\sim}{u})|_e^2 &= \langle |S_m(\underset{\sim}{u})|^2\rangle_M - |\langle S_m(\underset{\sim}{u})\rangle_M|^2 \\ &= |F(\underset{\sim}{u})|^2 \left(\langle |H_m(\underset{\sim}{u})|^2\rangle_M - |\langle H_m(\underset{\sim}{u})\rangle_M|^2\right). \end{aligned} \tag{3.27}$$

This has been called <u>long exposure subtraction</u> [Barakat and Nisenson 1983, cf. Nisenson et al 1983a]. Bates [1982b ϕ6.2] suggests that a suitably chosen weighting constant C, inserted into (3.27) to give

$$|F(\underset{\sim}{u})|_e^2 = \langle |S_m(\underset{\sim}{u})|^2\rangle_M - C\,|\langle S_m(\underset{\sim}{u})\rangle_M|^2 , \tag{3.28}$$

improves the compensation of the low spatial frequency excess of $\langle |H_m(\underset{\sim}{u})|^2\rangle_M$. C is determined empirically from observational experience. Bruck and Sodin [1980] propose a modified version of (3.28),

$$|F(\underset{\sim}{u})|_e^2 = \langle |S_m(\underset{\sim}{u})|^2\rangle_M - C(\underset{\sim}{u})\,|\langle S_m(\underset{\sim}{u})\rangle_M|^2 , \tag{3.29}$$

in which the weighting constant C is replaced by a polynomial of low, even order, chosen to minimise the effective width of $\mathbf{A}[f(\underset{\sim}{x})]_e = \mathbf{F}^{-1}[|F(\underset{\sim}{u})|_e^2]$. This criterion of choice stems from the suggestion made by Bates and Napier [1972] that errors in estimating the visibility of an image usually cause the effective width of the image to increase. Barakat and Nisenson [1983] compare the Welter-

Worden method (3.25, 3.27) with division by $\langle|H_{om}(\underset{\sim}{u})|^2\rangle_{Mo}$ from observations of a reference object (3.20). They show that the Welter-Worden method compensates inaccurately for the Labeyrie transfer function when the seeing is poor and conclude that division by a measured $\langle|H_{om}(\underset{\sim}{u})|^2\rangle_{Mo}$ is generally superior. Baba et al [1984] demonstrate the efficacy of Bruck and Sodin's [1980] method (3.29) in two-dimensional computer simulations.

Granrath [1984] suggests that the magnitude of $F(\underset{\sim}{u})$ can be obtained without having to compensate for an atmospheric transfer function by taking the maximum magnitude of $S(\underset{\sim}{u})$ at each spatial frequency across the ensemble of speckle images, forming

$$\begin{aligned} \max_m[\,|S_m(\underset{\sim}{u})|\,] &= |F(\underset{\sim}{u})|\ \max_m[H_m(\underset{\sim}{u})] \\ &= |F(\underset{\sim}{u})|\ H_G(\underset{\sim}{u}), \end{aligned} \tag{3.30}$$

instead of forming $\langle|S_m(\underset{\sim}{u})|^2\rangle_M$ as in speckle interferometry. The idea behind this is that $|H_m(\underset{\sim}{u})|$ is a random quantity which for each spatial frequency takes a value between zero and $H_D(\underset{\sim}{u})$, the diffraction-limited transfer function of the telescope. By taking the maximum magnitude at each spatial frequency of sufficiently many speckle images, the Granrath transfer function $H_G(\underset{\sim}{u})$ (3.30) should converge to $H_D(\underset{\sim}{u})$ [cf. Mariotti et al 1983 ϕ2.2.1]. Granrath [1984] assumes a linear probability density function for $|H_G(\underset{\sim}{u})|$ and then shows that $H_G(\underset{\sim}{u})$ can be used to form an unbiased estimate of $H_D(\underset{\sim}{u})$ whose variance decreases with increasing M faster than the variance of the average $\langle|H_m(\underset{\sim}{u})|^2\rangle_M$ does. However, I believe that the assumption of a uniform pdf is unrealistic. At spatial frequencies above the seeing limit but appreciably below the diffraction limit, for which the redundancy of a conventional aperture is high (ϕ1.6.5), the pdf of $|H_m(\underset{\sim}{u})|$ is strongly skewed towards zero. Figures 3.3a-d show frequency histograms of $|H_m(\underset{\sim}{u})|$ obtained at various spatial frequencies $|\underset{\sim}{u}|$ in a one-dimensional computer simulation. These frequency histograms, which are estimates of the pdf of $|H_m(\underset{\sim}{u})|$, show the unlikelihood of $|H_m(\underset{\sim}{u})|$ approaching $H_D(\underset{\sim}{u})$ for $|\underset{\sim}{u}|$ just above the seeing limit. This suggests that $H_G(\underset{\sim}{u})$ cannot be unbiased simply as Granrath [1984] suggests and that it would take an enormous number of speckle images for $H_G(\underset{\sim}{u})$ to converge usefully to $H_D(\underset{\sim}{u})$. Figures 3.3e-h show the Granrath transfer function for three

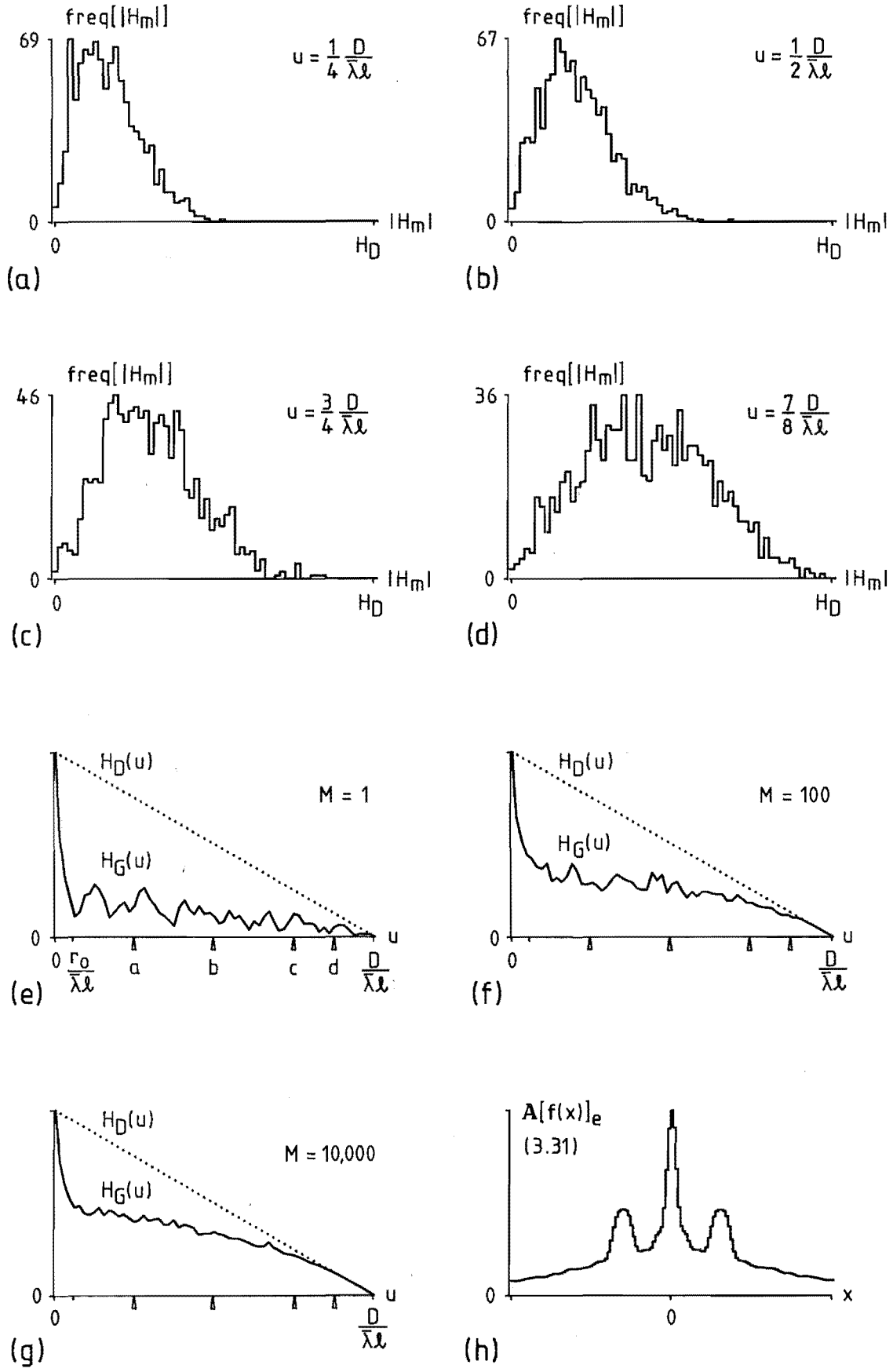

freq[|Hm|]
69
0
u = 1/4 D/λℓ
|Hm|
0
HD
(a)
freq[|Hm|]
67
0
u = 1/2 D/λℓ
|Hm|
0
HD
(b)
freq[|Hm|]
46
0
u = 3/4 D/λℓ
|Hm|
0
HD
(c)
freq[|Hm|]
36
0
u = 7/8 D/λℓ
|Hm|
0
HD
(d)
HD(u)
HG(u)
M = 1
0
u
0 r0/λℓ
a
b
c
d
D/λℓ
(e)
HD(u)
HG(u)
M = 100
0
u
0
D/λℓ
(f)
HD(u)
HG(u)
M = 10,000
0
u
0
D/λℓ
(g)
A[f(x)]e
(3.31)
0
x
0
(h)

Figure 3.3

values of M up to 10,000, demonstrating the slow convergence to $H_D(\underset{\sim}{u})$ just above the seeing limit. Even for ten thousand speckle images the convergence is poor (Figure 3.3g), leaving an appreciable fog (Figure 3.3h) in the estimated autocorrelation obtained from

$$\begin{aligned} \mathbf{A}[f(\underset{\sim}{x})]_e &= \mathbf{F}^{-1}[\max_m[\,|S_m(\underset{\sim}{u})|\,]^2] \\ &= \mathbf{F}^{-1}[\,|F(\underset{\sim}{u})|^2 H_G^{\,2}(\underset{\sim}{u})]\,, \end{aligned} \tag{3.31}$$

with M = 10,000, of the object shown in Figure 3.2a.

3.3.3 PHOTON LIMITING AND SIGNAL-TO-NOISE RATIO

The speckle image $s_m(\underset{\sim}{x})$ is a random quantity because random processes operate in its formation. At visible wavelengths there are two main sources of randomness in $s_m(\underset{\sim}{x})$: the random turbulence of the Earth's atmosphere, and the random detection of photons of light [Dainty and Shaw 1974 ϕ1]. Additive noise from sky background radiation and detector noise is a significant third source of randomness at infrared wavelengths [Sibille et al 1979] but is generally not considered important at visible wavelengths with modern

Figure 3.3 (facing page): One-dimensional computer simulation of Granrath's [1984] maximum-magnitude speckle interferometry proposal, which forms $\max_m[\,|S_m(\underset{\sim}{u})|\,]$ at each $\underset{\sim}{u}$. These simulations are of an unresolved object so that $S_m(\underset{\sim}{u}) = H_m(\underset{\sim}{u})$.

(a) - (d) Frequency histograms of the value of $|H_m(\underset{\sim}{u})|$ at the four spatial frequencies shown in (e), each compiled from 1000 speckle images.

(e) - (g) The Granrath transfer function $H_G(\underset{\sim}{u})$ (3.30) from M speckle images, compared to the diffraction-limited transfer function, for (e) M = 1, (f) M = 100, (g) M = 10,000.

(h) The Granrath estimate (3.31) of the autocorrelation of the object whose true image is shown in Figure 3.2a (cf. Figures 3.2g,l).

optical detectors [Dainty 1984 §7.3]. The turbulence of the atmosphere makes the classical intensity of the speckle image, i.e. its expected intensity in the limiting case of a very bright object, random (§2.5). The fact that light is composed of photons introduces additional randomness into the speckle image as well as causing its value to be quantised into discrete levels. Speckle images in which the photon effects are significant are termed photon-limited.

The power spectrum or autocorrelation averages $S_{LA}(\underset{\sim}{u})$, $s_{LA}(\underset{\sim}{x})$ formed by speckle interferometry (3.12, 3.14) are averages over an ensemble of M speckle images. They are themselves random quantities since the $s_m(\underset{\sim}{x})$ are random and M is always finite in practice. Their randomness introduces uncertainty into any quantity estimated from speckle interferometry. Dainty [1984 §7.3] defines the signal-to-noise ratio of an uncertain estimate Q of a quantity as

$$\text{SNR} = \frac{\text{expected value of Q}}{\text{standard deviation of Q}} = \frac{E[Q]}{\sqrt{\text{Var}[Q]}} . \tag{3.32}$$

From this, Dainty [1984] provides an extensive review of the signal-to-noise ratios of various quantities that can be estimated using speckle interferometry. This subsection quotes results pertaining to one such quantity, the averaged power spectrum $S_{LA}(\underset{\sim}{u}) = \langle |S_m(\underset{\sim}{u})|^2 \rangle_M$.

The discussion is assisted by introducing the following notation. Let $\Phi_f(\underset{\sim}{u})$ be the power spectrum of the object,

$$\Phi_f(\underset{\sim}{u}) = |F(\underset{\sim}{u})|^2 , \tag{3.33}$$

and let $\Phi_s(\underset{\sim}{u})$ be the expected value of the power spectrum of non-photon-limited speckle images of the object,

$$\Phi_s(\underset{\sim}{u}) = \langle |S_m(\underset{\sim}{u})|^2 \rangle . \tag{3.34}$$

Speckle interferometry in the absence of photon limiting (3.12) estimates $\Phi_s(\underset{\sim}{u})$. When $D >> r_o$,

$$\Phi_s(\underset{\sim}{u}) = \Phi_f(\underset{\sim}{u}) \langle |H_m(\underset{\sim}{u})|^2 \rangle$$

$$\approx A^2 \Phi_f(\underset{\sim}{u}) \left(|\hat{H}_{LE}(\underset{\sim}{u})|^2 + \frac{1}{N_{sp}} \hat{H}_D(\underset{\sim}{u}) \right) \qquad (3.35)$$

(3.13). Introducing the normalised power spectra

$$\hat{\Phi}_s(\underset{\sim}{u}) = \Phi_s(\underset{\sim}{u}) / \Phi_s(\underset{\sim}{0}),$$

$$\hat{\Phi}_f(\underset{\sim}{u}) = \Phi_f(\underset{\sim}{u}) / \Phi_f(\underset{\sim}{0}) \qquad (3.36)$$

which have unit value at the origin of Fourier space, (3.35) becomes

$$\hat{\Phi}_s(\underset{\sim}{u}) \approx \hat{\Phi}_f(\underset{\sim}{u}) \left(|\hat{H}_{LE}(\underset{\sim}{u})|^2 + \frac{1}{N_{sp}} \hat{H}_D(\underset{\sim}{u}) \right) \qquad (3.37)$$

where N_{sp} is the average number of speckles per speckle image (2.14). For all $|\underset{\sim}{u}|$ above the seeing limit, $\hat{H}_{LE}(\underset{\sim}{u}) \approx 0$ and so

$$\hat{\Phi}_s(\underset{\sim}{u}) \approx \frac{1}{N_{sp}} \hat{\Phi}_f(\underset{\sim}{u}) \hat{H}_D(\underset{\sim}{u}). \qquad (3.38)$$

Let $\bar{N}_p$ be the average number of photons detected per speckle image and $\bar{n}_p$ be the average number of photons detected per speckle,

$$\bar{n}_p = \frac{\bar{N}_p}{\bar{N}_{sp}} . \qquad (3.39)$$

$\bar{N}_p$ and N_{sp} are both proportional to D^K where K is the dimensionality of the aperture (K = 2 for a circular aperture). $\bar{n}_p$ is independent of D but is proportional to r_o^K (cf. 2.14) when $D >> r_o$. Let $s'_m(\underset{\sim}{x}) \leftrightarrow S'_m(\underset{\sim}{u})$ denote a photon-limited speckle image with classical intensity $s_m(\underset{\sim}{x})$.

When speckle images are photon-limited with $\bar{N}_p$ photons on average per speckle image, and the response of the optical detector to a single photon impact is modelled as a delta function in image space, Goodman and Belsher [1976] show that

$$\langle |S'_m(\underset{\sim}{u})|^2 \rangle = \bar{N}_p^2 \hat{\Phi}_s(\underset{\sim}{u}) + \bar{N}_p$$

$$= \bar{N}_p^2 \left(\Phi_s(\underset{\sim}{u}) + 1 / \bar{N}_p \right). \qquad (3.40a)$$

If the detector response to a photon impact is described by a point

spread function $h_d(\underset{\sim}{x}) \leftrightarrow H_d(\underset{\sim}{u})$ (this is not to be confused with the diffraction-limited psf $h_D(\underset{\sim}{x})$ and transfer function $H_D(\underset{\sim}{u})$), then (3.40a) becomes

$$\langle |S'_m(\underset{\sim}{u})|^2 \rangle = (\bar{N}_p^2 \, \hat{\Phi}_s(\underset{\sim}{u}) + \bar{N}_p) \, |H_d(\underset{\sim}{u})|^2 \qquad (3.40b)$$

[cf. Nisenson and Papaliolios 1983 §III, Dainty 1984 §7.3.1]. The estimate of $\Phi_s(\underset{\sim}{u})$ produced by speckle interferometry thus sits on top of a pedestal of height $\bar{N}_p$, whose height relative to the information-bearing term $\bar{N}_p^2 \, \hat{\Phi}_s(\underset{\sim}{u})$ in (3.40) is inversely proportional to $\bar{N}_p$. This pedestal is called the photon bias. If $\bar{N}_p$ and $|H_d(\underset{\sim}{u})|$ are known, the photon bias can readily be subtracted from $S'_{LA}(\underset{\sim}{u}) = \langle |S'_m(\underset{\sim}{u})|^2 \rangle_M$ to provide an unbiased estimate of $\hat{\Phi}_s(\underset{\sim}{u})$. If $\bar{N}_p$ is unknown but $H_d(\underset{\sim}{u})$ is wider in Fourier space than is $\hat{\Phi}_s(\underset{\sim}{u})$, then $\bar{N}_p$ can be estimated from $S'_{LA}(\underset{\sim}{u})$ by inspecting $S'_{LA}(\underset{\sim}{u})$ at spatial frequencies above the seeing limit $|\underset{\sim}{u}| = D/\bar{\lambda}\ell$, for which

$$\langle |S'_m(\underset{\sim}{u})|^2 \rangle = \bar{N}_p \, |H_d(\underset{\sim}{u})|^2 \qquad (3.41)$$

[cf. Hege et al 1982b §6]. $H_d(\underset{\sim}{u})$ is wider than $\hat{\Phi}_s(\underset{\sim}{u})$ when the detector resolution exceeds the telescope resolution, i.e. when the detector pixel size is smaller than the speckle size.

Aime et al [1985] use two detectors to form photon-limited speckle images $s'_{1m}(\underset{\sim}{x})$ and $s'_{2m}(\underset{\sim}{x})$ having the same classical intensity but statistically independent photon noises. This arrangement is equivalent to that of Hanbury Brown's [1974] intensity interferometer. The averaged cross-spectrum $\langle S'^*_{1m}(\underset{\sim}{u}) \, S'_{2m}(\underset{\sim}{u}) \rangle_M$ of $s'_{1m}(\underset{\sim}{x})$ and $s'_{2m}(\underset{\sim}{x})$ estimates $\hat{\Phi}_s(\underset{\sim}{u})$ with no photon bias.

Dainty [1984 §7.3.1] derives an approximate expression for the signal-to-noise ratio at a point in the speckle interferometry estimate of the power spectrum $\hat{\Phi}_s(\underset{\sim}{u})$ obtained by subtracting the photon bias from $\langle |S_m(\underset{\sim}{u})|^2 \rangle_M$ (3.40). The expression,

$$SNR \approx \frac{\bar{N}_p \hat{\Phi}_s(\underset{\sim}{u})}{1 + \bar{N}_p \hat{\Phi}_s(\underset{\sim}{u})} \tag{3.42a}$$

$$\approx \frac{\bar{n}_p \hat{\Phi}_f(\underset{\sim}{u}) \hat{H}_D(\underset{\sim}{u})}{1 + \bar{n}_p \hat{\Phi}_f(\underset{\sim}{u}) \hat{H}_D(\underset{\sim}{u})} , \tag{3.42b}$$

applies to the estimate of $\hat{\Phi}_s(\underset{\sim}{u})$ from a single speckle image at high spatial frequencies, say $|\underset{\sim}{u}| > 0.5\, D/\bar{\lambda}\ell$. The signal-to-noise ratio from an ensemble of M statistically independent speckle images increases proportionally to $\sqrt{M}$,

$$SNR_M = SNR.\sqrt{M} . \tag{3.43}$$

The limiting case of (3.42) for a very bright object, $\bar{n}_p \hat{\Phi}_f(\underset{\sim}{u}) \hat{H}_D(\underset{\sim}{u}) >> 1$, is

$$SNR \approx 1 . \tag{3.44}$$

The limiting case of (3.42) for a very dim object, $\bar{n}_p \hat{\Phi}_f(\underset{\sim}{u}) \hat{H}_D(\underset{\sim}{u}) << 1$, is

$$SNR \approx \bar{N}_p \hat{\Phi}_s(\underset{\sim}{u}) \tag{3.45a}$$

$$\approx \bar{n}_p \hat{\Phi}_f(\underset{\sim}{u}) \hat{H}_D(\underset{\sim}{u}). \tag{3.45b}$$

Equation (3.45) holds in practice for all fainter astronomical objects. Equations (3.42 - 3.45) have the following consequences:

(a) The signal-to-noise ratio for a single speckle image cannot exceed unity, the limit set by the randomness of atmospheric fluctuations. For bright astronomical objects the SNR approaches unity, and averaging over perhaps a few hundred speckle images raises SNR_M to a useful value [Nisenson and Papaliolios 1983]. For dim astronomical objects the photon noise greatly decreases the SNR per speckle image, with the consequence that many thousands of speckle images may be required to obtain a useful SNR_M under photon-limited conditions [Nisenson and Papaliolios 1983, cf. Dainty 1984 φ7.3.3].

(b) For dim objects (3.45), the SNR is proportional to $\bar{n}_p$ which is

proportional to r_o^K, i.e. the SNR is strongly dependent on the seeing. The SNR is independent of telescope diameter. However a larger telescope increases the number of independent points in the power spectrum in proportion to D^K.

3.4 SPECKLE IMAGING: POST-PROCESSORS TO SPECKLE INTERFEROMETRY

This section describes speckle imaging techniques which operate on the autocorrelation estimate or power spectrum estimate of the object generated by speckle interferometry. These techniques all effectively perform Fourier phase retrieval. Speckle holography and related techniques (ϕ3.4.1) exploit the properties of special forms of object to perform Fourier phase retrieval simply. Section 3.4.2 describes Fourier phase retrieval algorithms for objects of arbitrary form.

3.4.1 SPECKLE HOLOGRAPHY

Suppose that the object comprises two spatially separate parts,

$$f(\underset{\sim}{x}) = f_r(\underset{\sim}{x}) + f_u(\underset{\sim}{x}), \tag{3.46}$$

which are here called the reference object and the unknown object respectively. If the angular separation between the two parts is large enough relative to their spatial extents to satisfy the holographic separation condition (see Figure 3.4) [cf. Bates 1982b ϕ8.2], then the autocorrelation of $f(\underset{\sim}{x})$,

$$\mathbf{A}[f(\underset{\sim}{x})] = \mathbf{A}[f_r(\underset{\sim}{x})] + \mathbf{A}[f_u(\underset{\sim}{x})] + f_r(\underset{\sim}{x}) \otimes f_u(\underset{\sim}{x}) + f_u(\underset{\sim}{x}) \otimes f_r(\underset{\sim}{x}), \tag{3.47}$$

consists of three spatially separate parts (Figure 3.4b). The outer two of these parts, being the last two terms on the right hand side of (3.47), are the cross-correlations of the reference object and the unknown object,

$$f_r(\underset{\sim}{x}) \otimes f_u(\underset{\sim}{x}) = f_r(-\underset{\sim}{x}) \odot f_u(\underset{\sim}{x}) \quad \text{and} \tag{3.48a}$$

$$f_u(\underset{\sim}{x}) \otimes f_r(\underset{\sim}{x}) = f_r(\underset{\sim}{x}) \odot f_u(-\underset{\sim}{x}) \tag{3.48b}$$

(remember that $f_r(\underset{\sim}{x})$ and $f_u(\underset{\sim}{x})$ are real-valued). If the reference object is centrosymmetric (in the sense that $f_r(-\underset{\sim}{x}) = f_r(\underset{\sim}{x} + \underset{\sim}{b})$ for some constant $\underset{\sim}{b}$) and its form is known, then the form of the unknown object can be immediately recovered by deconvolving $f_r(\underset{\sim}{x})$ from (3.48a) or (3.48b) [cf. Bates 1971, Napier and Bates 1971]. If the reference object is unresolved, for example

$$f_r(\underset{\sim}{x}) = a\, \delta(\underset{\sim}{x} + \underset{\sim}{x}_r) \tag{3.49}$$

(i.e. a delta function of strength a at $\underset{\sim}{x} = \underset{\sim}{x}_r$), (3.48) becomes

$$f_r(\underset{\sim}{x}) \otimes f_u(\underset{\sim}{x}) = a\, f_u(\underset{\sim}{x} + \underset{\sim}{x}_r), \tag{3.50a}$$

$$f_u(\underset{\sim}{x}) \otimes f_r(\underset{\sim}{x}) = a\, f_u(-(\underset{\sim}{x} - \underset{\sim}{x}_r)). \tag{3.50b}$$

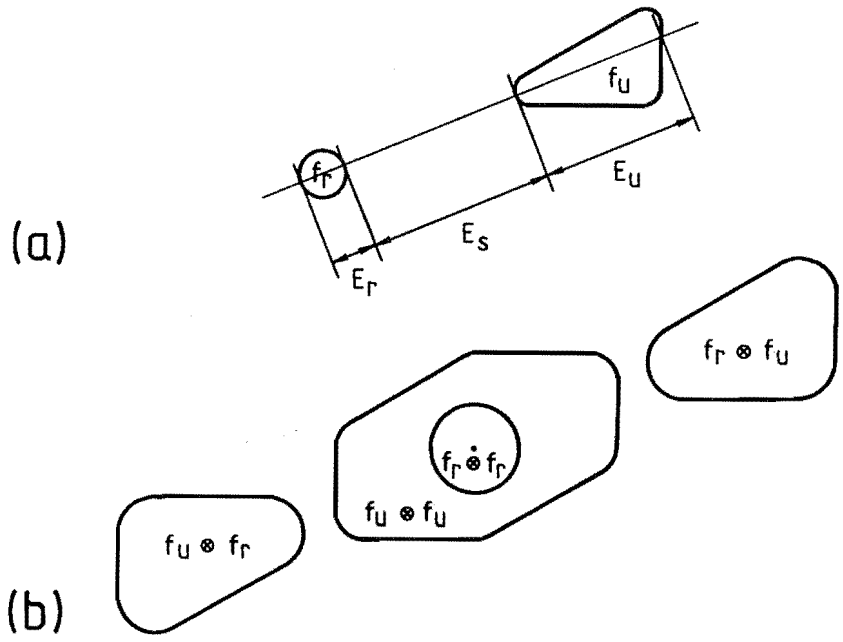


Figure 3.4: An object comprising two parts f_r and f_u satisfies the holographic separation condition if $E_s > \max[E_r, E_u]$ for all straight lines intersecting both parts (a). The autocorrelation of the object then consists of three spatially separate parts, two of these being cross-correlations of f_r and f_u (b).

The cross-correlation terms in (3.47) are then replicas of the unknown object, the first (3.50a) centred at $-\underset{\sim}{x}_r$ and the second (3.50b) its mirror image centred at $\underset{\sim}{x}_r$. The unresolved reference object acts like the reference wave in offset-reference holography [Goodman 1970].

The application of this principle to the estimated autocorrelations obtained from speckle interferometry is called speckle holography and was originally suggested by Bates et al [1973]. Gough and Bates [1974] report simulations of speckle holography in the optical laboratory. The efficacy of speckle holography in astronomical practice has been demonstrated by Weigelt and his colleagues [Weigelt 1978, Weigelt 1979, Ebersberger and Weigelt 1979, Weigelt 1980, Weigelt and Baier 1985, Weigelt et al 1985a]. Weigelt [1978] shows that an image of the entire object $f(\underset{\sim}{x})$ can be recovered if $f(\underset{\sim}{x})$ contains more than one part that can be used as a reference object for speckle holography.

Because of its demand for a reference object at an appreciable angular distance from the unknown object, the effects of nonisoplanatism (ϕ2.6) are of particular concern when applying speckle holography. The behaviour of the cross-correlations (3.50) is governed by the behaviour of the speckle cross-correlation function (2.28), which is described in Fourier space by (2.30). The first term on the right hand side of (2.30) is, ideally, eliminated when the Labeyrie transfer function (ϕ3.3.1) is compensated for in the normal process of speckle interferometry. This leaves the second term on the right hand side of (2.30) describing the spatially variant point spread function for the cross-correlation terms (3.50). Under increasing degrees of nonisoplanatism this psf becomes broader and its magnitude decreases, as is discussed in detail in Section 2.6. Partial isoplanatism can persist over angular separations of 10 arc seconds or more (ϕ2.6).

If an object containing an unresolved reference object does not satisfy the holographic separation criterion, then in the autocorrelation of the object the cross-correlations of the unknown object with the reference object overlap the central autocorrelations of each part of the object, and may even overlap each other

(cf. Figure 3.4b). The cross-correlation terms then cannot be easily identified by inspection. Additional information is required to separate an estimate of the unknown object from $\mathbf{A}[f(\underset{\sim}{x})]$.

Liu and Lohmann [1973] describe a technique here called large-field speckle holography in which the limited visibility phase information present in the long exposure image allows the overlapping cross-correlations to be distinguished from each other. The technique applies to objects having a reference part and several unknown parts or "islands" sufficiently far apart to be resolved from each other in a long exposure image. Also the cross-correlations between all of the "islands" must not overlap each other in the autocorrelation of the object. The spatial extent of such an object is necessarily large, hence "large field", and the comments on nonisoplanatism made above are especially pertinent. The particular islands in the autocorrelation of $f(\underset{\sim}{x})$ belonging to $f_r(\underset{\sim}{x}) \otimes f_u(\underset{\sim}{x})$ (3.50a), which represents a true image of the unknown object, are identified using the long exposure image. A mask $s_H(\underset{\sim}{x})$ is formed by hard-limiting the long exposure image, and the mask is placed over the estimated autocorrelation from speckle interferometry with an offset $\underset{\sim}{x}'$ chosen so that the cross-correlation

$$\int s_H(\underset{\sim}{x} + \underset{\sim}{x}')\, \mathbf{A}[f(\underset{\sim}{x})]_e \, .d\underset{\sim}{x} \tag{3.51}$$

is maximised. The product

$$s_H(\underset{\sim}{s} + \underset{\sim}{x}')\, \mathbf{A}[f(\underset{\sim}{x})]_e \tag{3.52}$$

is then, ideally, a true image of the unknown object together with a portion of the central autocorrelation terms of (3.50) in the place of the reference object. Liu and Lohmann [1973] show optical laboratory simulations of this procedure. They suggest that the chance of (3.52) being an "incorrect" estimate of the true image, due to the maximum of (3.51) occurring at an "incorrect" mask position, can be reduced for objects with more than one unresolved part by performing the process with different unresolved parts used as the reference object and requiring that the estimates (3.52) so obtained be similar.

If the brightness of the reference object $f_r(\underset{\sim}{x}) = a\, \delta(\underset{\sim}{x} - \underset{\sim}{x}_r)$

varies, the brightness of the cross-correlation terms of $f_r(\underset{\sim}{x})$ with $f_u(\underset{\sim}{x})$ in (3.47) varies in sympathy whereas the brightness of the $\mathbf{A}[f_u(\underset{\sim}{x})]$ term does not. Autocorrelations of the object formed with the reference object at different brightnesses, say

$$f_{r1}(\underset{\sim}{x}) = a_1\ \delta(\underset{\sim}{x} - \underset{\sim}{x}_r), \tag{3.53a}$$

$$f_{r2}(\underset{\sim}{x}) = a_2\ \delta(\underset{\sim}{x} - \underset{\sim}{x}_r), \tag{3.53b}$$

$$\mathbf{A}[f_1(\underset{\sim}{x})] = a_1^2\,\delta(\underset{\sim}{x}) + \mathbf{A}[f_u(\underset{\sim}{x})] + a_1 f_u(\underset{\sim}{x} + \underset{\sim}{x}_r) + a_1 f_u(-(\underset{\sim}{x} - \underset{\sim}{x}_r)), \tag{3.54a}$$

$$\mathbf{A}[f_2(\underset{\sim}{x})] = a_2^2\,\delta(\underset{\sim}{x}) + \mathbf{A}[f_u(\underset{\sim}{x})] + a_2 f_u(\underset{\sim}{x} + \underset{\sim}{x}_r) + a_2 f_u(-(\underset{\sim}{x} - \underset{\sim}{x}_r)), \tag{3.54b}$$

can then be subtracted to yield a difference

$$\mathbf{A}[f_1(\underset{\sim}{x})] - \mathbf{A}[f_2(\underset{\sim}{x})] = (a_1^2 - a_2^2)\,\delta(\underset{\sim}{x}) + (a_1 - a_2)\,[f_u(\underset{\sim}{x} + \underset{\sim}{x}_r) + f_u(-(\underset{\sim}{x} - \underset{\sim}{x}_r))] \tag{3.55}$$

in which the central autocorrelation term $\mathbf{A}[f_u(\underset{\sim}{x})]$ has cancelled out. The holographic separation condition is thus relaxed to the requirement $E_s > 0$ (cf. Figure 3.4a) for each cross-correlation to be spatially separate from the other terms in $\mathbf{A}[f(\underset{\sim}{x})]$. Rogers [1979] (see also Bates and Milner [1978 ϕ4]) makes this proposal for speckle imaging of star clusters containing a variable star. Rogers notes the similarity of the idea to the isomorphous replacement method of crystal structure determination in X-ray crystallography [Buerger 1959 ϕ8]. Lohmann and Weigelt [1975] and Weigelt [1975] (see also Lohmann and Weigelt [1977]) make the same suggestion as an extension of large-field speckle holography. Here, since the reference object creates a separate "speckle cloud" from the rest of the object in each speckle image, the brightness variation of the reference object can be introduced synthetically by appropriately masking its speckle cloud in each speckle image before performing speckle interferometry with the speckle images. Weigelt [1975] presents an optical simulation of this process, including an experiment with partially isoplanatic conditions.

If the reference object is much brighter than the unknown object in (3.46), then the first, third, and fourth terms on the right

hand side of (3.47) dominate $\mathbf{A}[f_u(\underset{\sim}{x})]$. All bright features of $\mathbf{A}[f(\underset{\sim}{x})]$ away from its origin are likely to belong to the cross-correlations of $f_r(\underset{\sim}{x})$ and $f_u(\underset{\sim}{x})$. Iterative techniques analogous to the heavy-atom method in X-ray crystallography [Buerger 1959 ϕ8] can then be used to extract and refine an estimate of $f_r(\underset{\sim}{x}) \otimes f_u(\underset{\sim}{x})$ or $f_u(\underset{\sim}{x}) \otimes f_r(\underset{\sim}{x})$ (3.48, 3.50) [Baldwin and Warner 1978a, Baldwin and Warner 1978b].

If the object consists of a collection of unresolved parts (e.g. a star cluster),

$$f(\underset{\sim}{x}) = \sum_{n=1}^{N} a_n \, \delta(\underset{\sim}{x} - \underset{\sim}{x}_n), \tag{3.56}$$

and no two of the parts have the same vector separation, then the only parts of the object's autocorrelation

$$\mathbf{A}[f(\underset{\sim}{x})] = \sum_{n'=1}^{N} \sum_{n''=1}^{N} a_{n'} a_{n''} \, \delta(\underset{\sim}{x} + \underset{\sim}{x}_{n'} - \underset{\sim}{x}_{n''}) \tag{3.57}$$

which overlap are the N terms ($\underset{\sim}{x}_{n'} = \underset{\sim}{x}_{n''}$) at the origin. The form of $f(\underset{\sim}{x})$ can then be recovered unambiguously from $\mathbf{A}[f(\underset{\sim}{x})]$ without the need for an additional reference object as shown by Christou [1981] (see also Fienup [1981b ϕ6] and Fienup et al [1982 Examples 5,6]). The first step is to determine N. This is possible given $\mathbf{A}[f(\underset{\sim}{x})]$ since P, the number of spatially separate parts in $\mathbf{A}[f(\underset{\sim}{x})]$, equals $N^2 - N + 1$ and hence

$$N = 1 + \frac{\sqrt{4P - 3}}{2} \, . \tag{3.58}$$

The second step is to determine the support of the object, i.e. the relative positions of its parts. This procedure is illustrated in Figure 3.5. The procedure yields the support of either $f(\underset{\sim}{x})$ or its mirror image $f(-\underset{\sim}{x})$, which cannot be distinguished from each other if only $\mathbf{A}[f(\underset{\sim}{x})]$ is given. The third step is to find the brightnesses a_n of the parts of the object (3.56) by using the support found in the second step as a mask to select replicas of $f(\underset{\sim}{x})$ or $f(-\underset{\sim}{x})$ from $\mathbf{A}[f(\underset{\sim}{x})]$. At least two different replicas must be selected since one part of each replica lies at the origin of $\mathbf{A}[f(\underset{\sim}{x})]$ and is contaminated by parts of all of the other replicas. The success of this algorithm when applied to the autocorrelation estimate $\mathbf{A}[f(\underset{\sim}{x})]_e$ obtained from speckle interferometry depends on all of the P distinct parts of

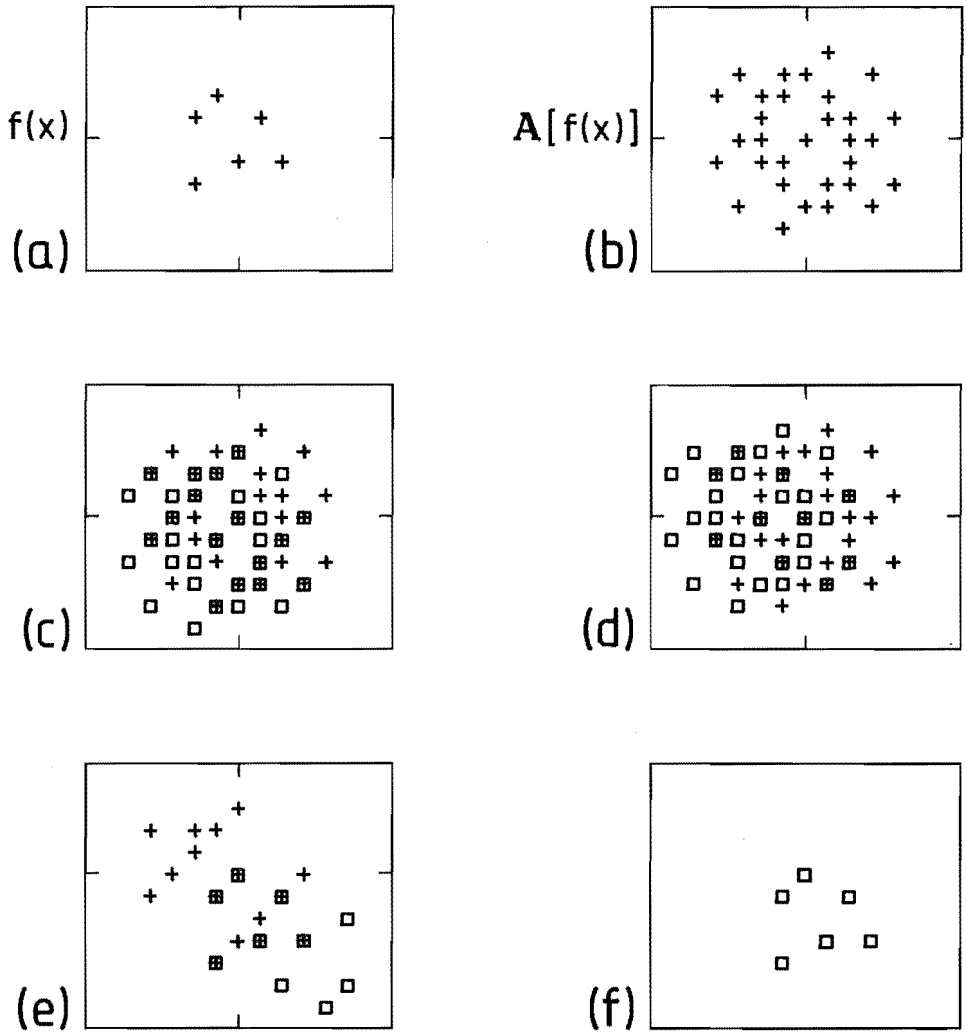


Figure 3.5: Algorithm for obtaining the support of an object $f(\underset{\sim}{x})$, consisting of unresolved parts with no repeated vector separations, from its autocorrelation $\mathbf{A}[f(\underset{\sim}{x})]$.

(a) The object, comprising N = 6 parts in this example.

(b) The object's autocorrelation, comprising $N^2 + N - 1$ parts.

The support of $f(\underset{\sim}{x})$ or $f(-\underset{\sim}{x})$ is deduced in the following steps:

(c) Translate (b) to bring each of its non-origin parts to the origin (□), and form the intersection of this with (b) (⊞).

(d) Repeat (c) with a different translation of (b).

(e) Translate the intersection found in (d) (□) so that it intersects with the intersection found in (c) (+) at N parts (⊞).

(f) The intersection found in (e) is the support of either $f(\underset{\sim}{x})$ or $f(-\underset{\sim}{x})$.

$\mathbf{A}[f(\underset{\sim}{x})]$ being recognisable in $\mathbf{A}[f(\underset{\sim}{x})]_e$. The brightness of each part of $f(\underset{\sim}{x})$ must not be so small that parts of $\mathbf{A}[f(\underset{\sim}{x})]$ become indistinguishable in noise, and the vector separations between the parts of $f(\underset{\sim}{x})$ must differ by at least the diffraction-limited resolution of the telescope. Christou's [1981] algorithm is a member of a more general set of algorithms described by Fienup [1981b], Fienup et al [1982], and Fienup [1983] for recovering an object from its autocorrelation when the holographic separation condition is not fulfilled.

3.4.2 FOURIER PHASE RETRIEVAL

The determination of $F(\underset{\sim}{u})$, the Fourier transform of a quantity of interest $f(\underset{\sim}{x})$, when the phase of $F(\underset{\sim}{u})$ is only partly known or is completely unknown, is called the Fourier phase problem [Bates and Fright 1984, Fright 1984]. The Fourier phase problem occurs in many situations in physical science where phase$[F(\underset{\sim}{u})]$ is inaccurately measured or is unmeasurable, for example in optical astronomy (Chapter 2), radio astronomy, electron microscopy, and X-ray crystallography [Bates and Fright 1984 §§3,4]. The process of inferring phase$[F(\underset{\sim}{u})]$ from $|F(\underset{\sim}{u})|$ and whatever else is known about $F(\underset{\sim}{u})$ or $f(\underset{\sim}{x})$ is here called Fourier phase retrieval. In optical astronomical speckle imaging, Fourier phase retrieval is invoked to estimate the object visibility phase from the visibility magnitude estimate $|F(\underset{\sim}{u})|_e$ generated by speckle interferometry, and hence form an estimate of the true image. Fourier phase retrieval in speckle imaging is reviewed by Bates [1982b §7] and Dainty [1984 §§7.4.1, 7.4.3]. This subsection briefly describes three Fourier phase retrieval algorithms that have been applied in speckle imaging.

In a completely general case, given only $|F(\underset{\sim}{x})|$, the Fourier phase problem has no unique solution because an arbitrary phase$[F(\underset{\sim}{u})]$ can be assigned to any $|F(\underset{\sim}{u})|$. However, additional knowledge of $F(\underset{\sim}{u})$ or $f(\underset{\sim}{x})$ places constraints on the phase$[F(\underset{\sim}{u})]$ that can be associated with a given $|F(\underset{\sim}{u})|$. The constraints become stronger the more additional knowledge there is. Firstly, the extent of the object $f(\underset{\sim}{x})$ is always effectively finite in practice, either because the source is of finite size or because it is observed with an imaging instrument

having a finite field of view. Secondly, $f(\underset{\sim}{x})$ is necessarily real and non-negative in optical astronomy because the corresponding source (§2.2) is spatially incoherent. These constraints on $f(\underset{\sim}{x})$ are sufficient to ensure an almost-always unique solution to the Fourier phase problem in dimensionalities of $f(\underset{\sim}{x})$ greater than one [e.g. Bruck and Sodin 1979, Bates 1982a, Bates 1984, Fright 1984 §3, Nieto-Vesperinas and Dainty 1984; cf. Dainty and Fiddy 1984, van Toorn et al 1984] (see Bates [1984] and Bates and Tan [1985] for results on complex $f(\underset{\sim}{x})$). There exists vigorous controversy over the question of uniqueness in two or more dimensions. This subsection follows the view of Bates and Fright [1984 §5], that the special forms of $f(\underset{\sim}{x})$ for which the solution is non-unique [e.g. Fienup 1983 §3C] possess special symmetries which never occur with actual measurements contaminated by inevitable noise. In the discussion which follows, the position vectors $\underset{\sim}{x}$ and $\underset{\sim}{u}$ are understood to have a dimensionality of two or higher (this implies two dimensions in the speckle imaging context), so that the statement of uniqueness made above applies below.

When given only $|F(\underset{\sim}{u})|$ and the knowledge that $f(\underset{\sim}{x})$ is of finite extent and non-negative, there are two <u>trivial characteristics</u> [cf. Bates 1982a §1] of phase$[F(\underset{\sim}{u})]$ which are unrecoverable. These are described by the expressions

$$\text{phase}[F(\underset{\sim}{u})] \sim \text{phase}[F(\underset{\sim}{u})] + \underset{\sim}{b}.\underset{\sim}{u} \ , \tag{3.59a}$$

$$\text{phase}[F(\underset{\sim}{u})] \sim -\text{phase}[F(\underset{\sim}{u})] \tag{3.59b}$$

where "~" is read as "is indistinguishable from" or "is not preferred to" and $\underset{\sim}{b}$ is an arbitrary vector constant. The corresponding effects on the reconstructed object $f(\underset{\sim}{x})$ are

$$f(\underset{\sim}{x}) \sim f(\underset{\sim}{x} - \underset{\sim}{b}), \tag{3.60a}$$

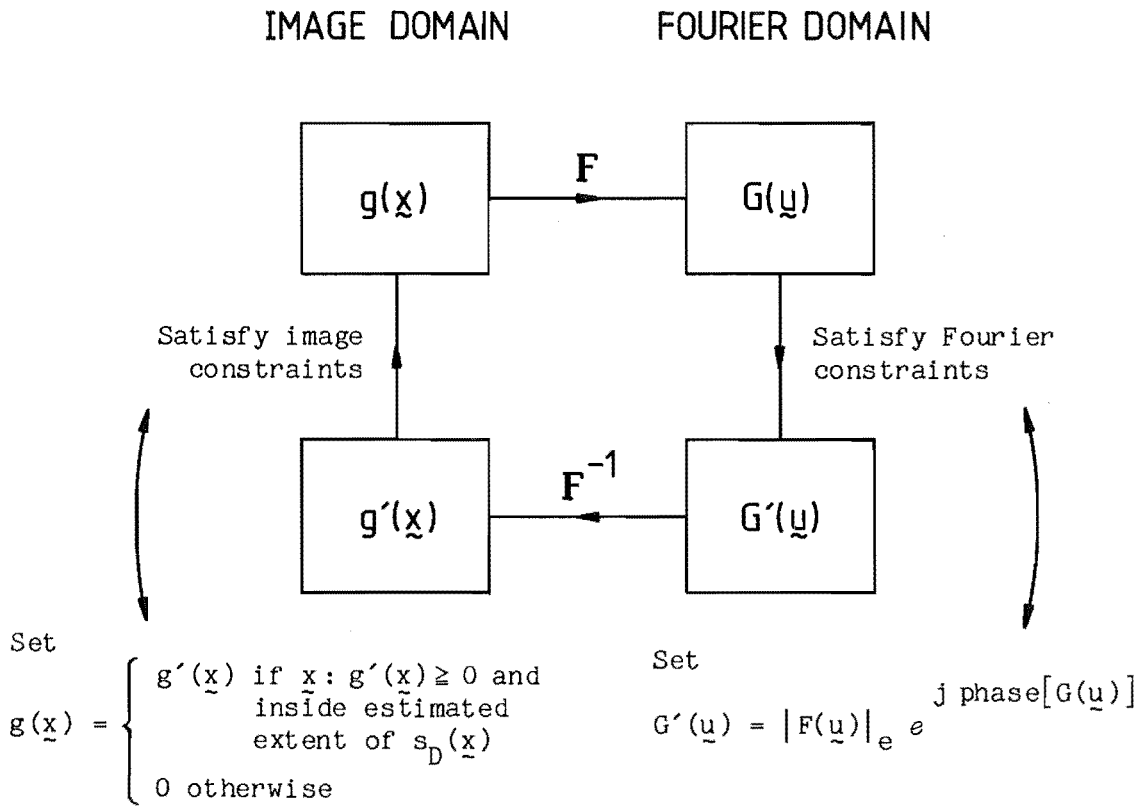
$$f(\underset{\sim}{x}) \sim f(-\underset{\sim}{x}), \tag{3.60b}$$

the first trivial characteristic (3.59a, 3.60a) being a translation of $f(\underset{\sim}{x})$ in image space and the second (3.59b, 3.60b) being a substitution of $f(\underset{\sim}{x})$ by its mirror image $f(-\underset{\sim}{x})$. Neither of these effects alter the image-form of $f(\underset{\sim}{x})$ (3.11), which is why the characteristics (3.59) are called trivial. The inability to distinguish between $f(\underset{\sim}{x})$ and its

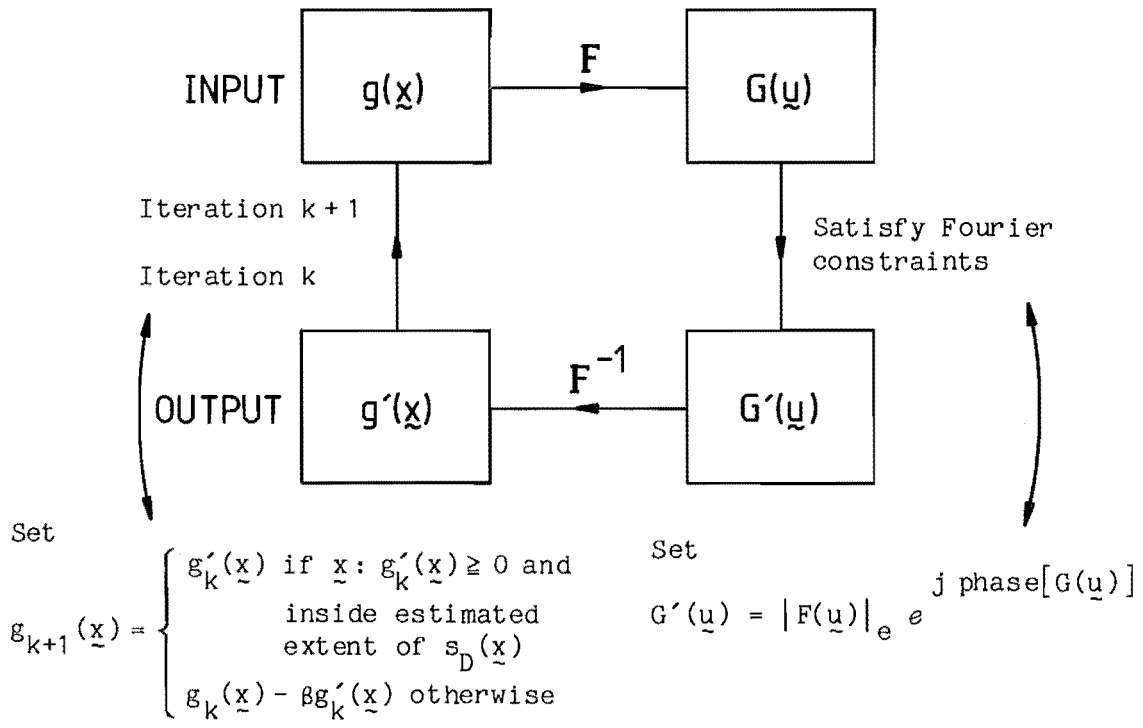
mirror image has already arisen several times in the discussion of speckle holography in Section 3.4.1.

Fienup [1978, 1981a, 1982, 1984] describes iterative algorithms for Fourier phase retrieval which are here called Fienup's algorithms. These are described in the speckle imaging context by Fienup [1979] and Fienup and Feldkamp [1980]. The algorithms transform an estimate of the true image $s_D(\underline{x})$ back and forth between image space and Fourier space, applying known constraints in each space at each iteration. The Fourier space constraint is the known visibility magnitude estimate $|F(\underline{u})|_e$ (cf. 3.20) obtained from speckle interferometry. The image space constraints are the non-negativity of $s_D(\underline{x})$ and the estimated extent of $s_D(\underline{x})$ (which is estimated as half the extent of the autocorrelation estimate $\mathbf{A}[f(\underline{x})]_e \leftrightarrow |F(\underline{u})|_e^2$). One of Fienup's algorithms, the error-reduction algorithm [Fienup 1982 ϕII] (see Figure 3.6a), is an adaptation of the earlier Gerchberg-Saxton algorithm [Gerchberg and Saxton 1972] for Fourier phase retrieval in electron microscopy. The mean squared error in the estimate of $s_D(\underline{x})$ has been proven to never increase with iteration count in the error-reduction algorithm [Fienup 1982 ϕII], which is effectively a guarantee of convergence towards a correct solution. However the convergence is usually extremely slow in practice [Fienup 1982]. The hybrid input-output algorithm (see Figure 3.6b) is a development of the error-reduction algorithm that does not have guaranteed convergence but does converge rapidly [Fienup 1982 ϕϕV,VI]. The error-reduction and hybrid input-output algorithms complement each other and are usefully applied alternately during a single phase reconstruction [cf. Fienup 1982 ϕVII]. Since Fienup's algorithms are iterative and involve two Fourier transforms per iteration, they tend to be computationally expensive.

A direct (non-iterative) Fourier phase retrieval algorithm called crude phase estimation or CPE is described by Bates and Fright [1983] and Bates and Fright [1984 ϕϕ10,11] [cf. Won et al 1985]. This algorithm uses algebraic relationships between the magnitudes of pixels in an oversampled representation of $F(\underline{u})$ to estimate $\text{phase}[F(\underline{u})]$ in a recursive manner somewhat similar to that of the Knox-Thompson method (ϕ3.5.2). Although recent extensions of CPE are promising [Won et al 1985], the estimate of $\text{phase}[F(\underline{u})]$ produced by



(a) Error-reduction algorithm



(b) Hybrid input-output algorithm

Figure 3.6: Two of Fienup's algorithms [Fienup 1984].

CPE is generally not as good as that which Fienup's algorithms can ultimately provide. CPE has therefore been combined with Fienup's algorithms in a composite Fourier phase retrieval algorithm called the Canterbury algorithm [Bates and Fright 1983, Bates and Fright 1984 §§11-13, Fright 1984 §7]. The crude phase estimate from CPE is used as a starting phase for Fienup's algorithms (phase$[G'(\underset{\sim}{u})]$ in Figure 3.6), which considerably reduces the number of iterations and hence the computational effort required to reach a given accuracy of phase retrieval. For foggy images having low-contrast detail on a uniform background, the Canterbury algorithm applies a defogging procedure which attenuates the magnitude of $F(\underset{\sim}{u})$ at low spatial frequencies prior to the application of CPE and Fienup's algorithms, and thereby considerably improves the quality of phase retrieval.

Gull and Daniell [1978] describe an iterative Fourier phase retrieval algorithm derived from the maximum entropy method of image restoration [Narayan and Nityananda 198_]. Gull and Daniell's algorithm artificially constrains $f(\underset{\sim}{x})$ by generating an estimate of the true image in which the "entropy" $\int -s(\underset{\sim}{x}) \log s(\underset{\sim}{x}) \, .d\underset{\sim}{x}$ is maximised subject to an estimate of the uncertainty in the given data $|F(\underset{\sim}{u})|_e$. Nityananda and Narayan [1982] show that maximum entropy image reconstruction works best with objects possessing isolated peaks on a flat background. The radio astronomical phase retrieval example of Gull and Daniell [1978] is of this kind. Roddier and Roddier [1985] present an image reconstruction of the disc of Betelgeuse, generated by a composite of Gull and Daniell's [1978] algorithm and Fienup's algorithms, starting with $|F(\underset{\sim}{u})|_e$ obtained from a pupil-plane interferometer (§2.8). Their reconstruction bears a striking qualitative resemblance [Roddier 1986] to the reconstruction of Betelgeuse from the same data by Bates et al [198_] using the Canterbury algorithm.

A priori knowledge of the form of the object, obtained from the long exposure image and by inference from astrophysical reasoning and observational techniques such as spectroscopy, can considerably assist Fourier phase retrieval. Lohmann and Weigelt [1977] describe a simple method called phase flipping which is applicable when $f(\underset{\sim}{x})$ is known to be centrosymmetric. $F(\underset{\sim}{u})$ is the purely real, so that

phase[$F(\underset{\sim}{u})$] equals either 0 or π and can be estimated by inspection of $|F(\underset{\sim}{u})|$. The techniques of speckle holography (ϕ3.4.1) are simple Fourier phase retrieval algorithms which are applicable when $f(\underset{\sim}{x})$ is known to have certain special forms.

The speckle imaging techniques described in Sections 3.5 and 3.6 generate, either explicitly or implicitly, estimates of phase[$F(\underset{\sim}{u})$] which can be used in place of a randomly generated phase [cf. Fienup 1984 ϕ7] to start Fienup's algorithms. The phase estimate does not have to be very accurate to significantly reduce the convergence time of Fienup's algorithms or to resolve the mirror-image ambiguity of image reconstruction associated with the trivial characteristic (3.59b, 3.60b). Fienup's algorithms lend themselves well to the construction of composite speckle imaging schemes because they are easily adapted to make use of partial knowledge of $f(\underset{\sim}{x})$ or phase[$F(\underset{\sim}{u})$]. Several such composite schemes have been described [Cocke 1980, Bates and Fright 1982, Baba et al 1984] using Fienup's algorithms to tidy the image produced by one of the speckle imaging methods described in Sections 3.5 and 3.6.

3.5 SPECKLE IMAGING: FOURIER DOMAIN METHODS

This section describes speckle imaging techniques whose theoretical descriptions or practical implementations are set, as for speckle interferometry, primarily in the Fourier domain. The various methods which explicitly attempt to retain the Fourier phase information discarded by speckle interferometry fall into this category.

3.5.1 PHASE AVERAGING

The phase of the speckle visibility as modelled by (3.3f) is the sum of the phases of the object visibility and the speckle transfer function,

$$\text{phase}[S_m(\underline{u})] = \text{phase}[F(\underline{u})H_m(\underline{u})]$$
$$= \text{phase}[F(\underline{u})] + \text{phase}[H_m(\underline{u})], \qquad (3.61)$$

where the unwrapped values [cf. Tribolet 1977] rather than the principal values (modulo-2π) of the phases are considered. Phase$[H_m(\underline{u})]$ is random for $|\underline{u}|$ greater than the seeing limit (ϕ2.5). Averaging (3.61) over the ensemble of speckle images produces

$$\langle \text{phase}[S_m(\underline{u})] \rangle_M = \text{phase}[F(\underline{u})] + \langle \text{phase}[H_m(\underline{u})] \rangle_M , \qquad (3.62)$$

in which the expected value of $\langle \text{phase}[H_m(\underline{u})] \rangle_M$ is, for given seeing conditions, a deterministic quantity that can be measured or predicted, allowing phase$[F(\underline{u})]$ to be estimated from $\langle \text{phase}[S_m(\underline{u})] \rangle_M$. This procedure is called phase averaging. McGlamery [1971] suggested it soon after Labeyrie [1970] introduced speckle interferometry (O'Connor and Huang [1981] describe McGlamery's algorithm in detail), and it has been suggested again more recently by Cocke [1980]. Since only the principal value of phase$[S_m(\underline{u})]$ is directly measurable, the phase must be unwrapped before (3.62) can be applied. This is not a trivial task [O'Connor and Huang 1981]. For isotropic turbulence one has $\langle \text{phase}[H_m(\underline{u})] \rangle = 0$ [O'Connor and Huang 1981, Dainty 1984 ϕ7.4.6] and hence

$$\langle \text{phase}[S_m(\underline{u})] \rangle_M \approx \text{phase}[F(\underline{u})]. \qquad (3.63)$$

The variance of $\langle \text{phase}[S_m(\underline{u})] \rangle_M$ increases with increasing $|\underline{u}|$ and is large where $|F(\underline{u})|$ is small, and decreases with increasing ensemble size in proportion to 1/M [O'Connor and Huang 1981, Dainty 1984 ϕ7.4.6].

The magnitude of $F(\underline{u})$ is estimated by McGlamery [1971] [cf. O'Connor and Huang 1981] and Cocke [1980] (and also by Mertz [1979], see later in this subsection) by averaging $|S_m(\underline{u})|$ to give

$$\langle |S_m(\underline{u})| \rangle_M = |F(\underline{u})| \, \langle |H_m(\underline{u})| \rangle_M . \qquad (3.64)$$

The transfer function $\langle |H_m(\underline{u})| \rangle_M$ must be estimated to allow $|F(\underline{u})|$ to be recovered from (3.64). Equation (3.64) is slightly different

from the procedure of speckle interferometry in which $|S_m(\underset{\sim}{u})|^2$ is averaged (3.12).

Cocke [1980] and O'Connor and Huang [1981] compare the performance of phase averaging with the Knox-Thompson method (§3.5.2) in computer simulations. They both find that the Knox-Thompson method gives slightly better estimates of the true image than phase averaging in the absence of photon noise. Cocke [1980] obtains better estimates from phase averaging than from the Knox-Thompson method (performed without correction of photon bias [Nisenson and Papaliolios 1983]) under photon-limited conditions.

Mertz [1979, 1984] describes an algorithm which improves the robustness of phase unwrapping by comparing or "tracking" phase$[S_m(\underset{\sim}{u})]$ at each spatial frequency through successive speckle images. The algorithm applies to speckle images whose exposure intervals are close together or even overlapping so that consecutive speckle images are statistically highly dependent and have phases that are well correlated with each other. Itoh and Ohtsuka [1983] describe a maximum likelihood algorithm for estimating phase$[F(\underset{\sim}{u})]$ from {phase$[S_m(\underset{\sim}{u})]$} which does not require the phase$[S_m(\underset{\sim}{u})]$ to be unwrapped. Aime et al [1975] propose phase averaging with speckle visibilities obtained from a pupil plane interferometer (§2.8).

3.5.2 PHASE GRADIENT AVERAGING; KNOX-THOMPSON METHOD

Phase averaging (§3.5.1) has the disadvantages that the speckle visibility phase must be unwrapped before being averaged and that the variance of the estimate of the object visibility phase is large at high spatial frequencies [Dainty 1984 §7.4.6]. An alternative approach is to average the gradient [Kreyszig 1979 §8.8] of the speckle visibility phase, forming

$$\langle \nabla \text{phase}[S_m(\underset{\sim}{u})] \rangle_M = \nabla \text{phase}[F(\underset{\sim}{u})] + \langle \nabla \text{phase}[H_m(\underset{\sim}{u})] \rangle_M \qquad (3.65a)$$

$$\approx \nabla \text{phase}[F(\underset{\sim}{u})] \qquad (3.65b)$$

[cf. Aitken et al 1985]. A discrete approximation to (3.65) is to average phase differences in $S_m(\underset{\sim}{u})$ over a finite frequency increment

$\Delta\underset{\sim}{u}$ in Fourier space [cf. Knox and Thompson 1974]. The phase gradient, or phase difference across a sufficiently small interval $|\Delta\underset{\sim}{u}|$, does not suffer from ambiguities due to the 2π principal range of phase$[S_m(\underset{\sim}{u})]$ because phase$[H_m(\underset{\sim}{u})]$ has a nonzero correlation length in Fourier space. Phase$[F(\underset{\sim}{u})]$ is evaluated by integrating the phase gradient or summing phase differences, starting at the origin of Fourier space where the real-valuedness of the object ensures that

$$\text{phase}[F(\underset{\sim}{0})] = 0 \ . \tag{3.66}$$

The best known speckle imaging technique employing the phase gradient is the Knox-Thompson method introduced by Knox and Thompson [1974]. This subsection describes the Knox-Thompson method in detail and briefly describes some other methods which are developments of the Knox-Thompson method or are otherwise connected with the phase gradient.

The Knox-Thompson method [Knox and Thompson 1974, Knox 1976] allows the object visibility phase to be estimated for an object of arbitrary form. The visibility phase estimate is combined with the visibility magnitude estimate from speckle interferometry to give an estimate of the true image. The Knox-Thompson method is distinct from an earlier proposal by Knox and Thompson [1973] which has been shown to be equivalent to long exposure imaging with auto-guiding (φ2.8) [Miller et al 1973, cf. Wildey 1973]. Bates [1982b φ8.4] notes that the Knox-Thompson method is closely related to an earlier technique of Jennison's [1958] for measuring the object visibility phase in radio astronomy.

In the Knox-Thompson method the ensemble autocorrelation of the speckle visibilities is formed at a displacement frequency $\Delta\underset{\sim}{u}$,

$$\langle S_m^*(\underset{\sim}{u})\, S_m(\underset{\sim}{u}+\Delta\underset{\sim}{u})\rangle_M = F^*(\underset{\sim}{u})\, F(\underset{\sim}{u}+\Delta\underset{\sim}{u})\, \langle H_m^*(\underset{\sim}{u})\, H_m(\underset{\sim}{u}+\Delta\underset{\sim}{u})\rangle_M \ . \tag{3.67}$$

This is a generalisation of speckle interferometry (3.12) which forms (3.67) with $\Delta\underset{\sim}{u} = \underset{\sim}{0}$. The Knox-Thompson transfer function $\langle H_m^*(\underset{\sim}{u})\, H_m(\underset{\sim}{u}+\Delta\underset{\sim}{u})\rangle$ is the expected value of the ensemble autocorrelation of the speckle transfer function, evaluated at $\Delta\underset{\sim}{u}$, and is shown by

Dainty [1984 §7.4.7] [cf. Knox and Thompson 1974 §III] to be given by

$$\langle H_m^*(\underset{\sim}{u})\, H_m(\underset{\sim}{u}+\Delta\underset{\sim}{u})\rangle \approx \langle |H_m(\underset{\sim}{u})|^2\rangle \; |H_A(\Delta\underset{\sim}{u}/2)|^2 \qquad (3.68)$$

(cf. 3.4, 2.8e) under the assumption of complex Gaussian statistics of the aperture plane field (§2.3.1). The magnitude of (3.68) is significantly greater than zero out to the diffraction limit of the telescope (cf. 3.13) provided $|\Delta\underset{\sim}{u}|$ is significantly less than the seeing limit $r_o/\bar{\lambda}\ell$ (§2.4). This means that information about $F^*(\underset{\sim}{u})\, F(\underset{\sim}{u}+\Delta\underset{\sim}{u})$ is available from (3.67) at frequencies out to the diffraction limit. For $\Delta\underset{\sim}{u} = \underset{\sim}{0}$, (3.67) is real and non-negative since it reduces to (3.12). For $\Delta\underset{\sim}{u} \neq \underset{\sim}{0}$, (3.67) is in general complex and conveys information about phase$[F(\underset{\sim}{u})]$. The phase of (3.67) provides an estimate of phase differences of $F(\underset{\sim}{u})$ across the spatial frequency increment $\Delta\underset{\sim}{u}$ since

$$\text{phase}[\langle S_m^*(\underset{\sim}{u})\, S_m(\underset{\sim}{u}+\Delta\underset{\sim}{u})\rangle_M] = \text{phase}[F(\underset{\sim}{u}+\Delta\underset{\sim}{u})] - \text{phase}[F(\underset{\sim}{u})] + \text{phase}[\langle H_m^*(\underset{\sim}{u})\, H_m(\underset{\sim}{u}+\Delta\underset{\sim}{u})\rangle_M]. \qquad (3.69)$$

For sufficiently small $|\Delta\underset{\sim}{u}|$ the value of (3.69) remains within the principal range $(-\pi;\pi]$. The third term on the right hand side of (3.69) is approximately zero for sufficiently large M [Knox 1976], since the right hand side of (3.68) is real-valued, and so

$$\text{phase}[\langle S_m^*(\underset{\sim}{u})\, S_m(\underset{\sim}{u}+\Delta\underset{\sim}{u})\rangle_M] \approx \text{phase}[F(\underset{\sim}{u}+\Delta\underset{\sim}{u})] - \text{phase}[F(\underset{\sim}{u})]. \qquad (3.70)$$

Phase$[F(\underset{\sim}{u})]$ is estimated at discrete locations in Fourier space by summing the phase differences (3.70) along paths leading away from the origin (cf. 3.66). In one dimension [Knox and Thompson 1974], the path of summation lies along the u-axis. Let

$$\phi_p = \text{phase}[F(p\Delta u)]_e \qquad (3.71a)$$

(with p an integer) be the estimate of phase$[F(\underset{\sim}{u})]$ to be solved for, and

$$\Delta\phi_p = \text{phase}[\langle S_m^*(p\Delta u)\, S_m((p+1)\Delta u)\rangle_M] = \phi_{p+1} - \phi_p \qquad (3.71b)$$

(cf. 3.70) be the given data from the Knox-Thompson averaging. Then, from (3.66) and (3.71),

$$\phi_0 = 0 \ , \tag{3.72a}$$

$$\phi_p = \sum_{n=0}^{p-1} \Delta\phi_n + \phi_0 \tag{3.72b}$$

for p = 1,2,... as far as the data permit (ideally, to the diffraction limit of the telescope). In two dimensions [Knox 1976] there is more than one possible path of summation to each point in Fourier space. Knox [1976] sums along paths parallel to the Cartesian axes of Fourier space, e.g.

$$\phi_{p,q} = \text{phase}\left[F(p\Delta u, q\Delta v)\right]_e \ , \tag{3.73a}$$

$$\Delta\phi_{p,q;p} = \text{phase}\left[\langle S_m^*(p\Delta u, q\Delta v)\, S_m((p+1)\Delta u, q\Delta v)\rangle_M\right], \tag{3.73b}$$

$$\Delta\phi_{p,q;q} = \text{phase}\left[\langle S_m^*(p\Delta u, q\Delta v)\, S_m(p\Delta u, (q+1)\Delta v)\rangle_M\right]; \tag{3.73c}$$

$$\phi_{0,0} = 0 \ , \tag{3.74a}$$

$$\phi_{p,0} = \sum_{n=0}^{p-1} \Delta\phi_{n,0;p} + \phi_{0,0} \ , \tag{3.74b}$$

$$\phi_{p,q} = \sum_{n=0}^{q-1} \Delta\phi_{p,n;q} + \phi_{p,0} \ . \tag{3.74c}$$

Equation (3.74a) gives the phase at the origin of Fourier space, (3.74b) gives the phase at points along the u-axis, and (3.74c) gives the phase at a rectangular grid of points lying on lines parallel to the v-axis. Baba et al [1984] sum along radial lines passing through the origin,

$$\phi_p = \text{phase}\left[F(p\Delta\underline{u})\right]_e \ , \tag{3.75a}$$

$$\Delta\phi_p = \text{phase}\left[\langle S_m(p\Delta\underline{u})\, S_m((p+1)\Delta\underline{u})\rangle_M\right]; \tag{3.75b}$$

$$\phi_0 = 0 \ , \tag{3.76a}$$

$$\phi_p = \sum_{n=0}^{p-1} \Delta\phi_n + \phi_0 \ . \tag{3.76b}$$

The phase estimate phase$[F(\underset{\sim}{u})]_e$ obtained in discrete form from (3.71 - 3.76) is combined with the magnitude estimate $|F(\underset{\sim}{u})|_e$ from speckle interferometry (3.20) to give an estimate

$$F(\underset{\sim}{u})_e = |F(\underset{\sim}{u})|_e \, e^{j \, \text{phase}[F(\underset{\sim}{u})]_e} \tag{3.77}$$

of $F(\underset{\sim}{u})$ within the diffraction limit of the telescope. The usual prescription for the Knox-Thompson method [e.g. Nisenson et al 1983b] then forms an estimate of the true image by inverse Fourier transforming (3.77), giving

$$s_{KT}(\underset{\sim}{x}) = \mathbf{F}^{-1}[F(\underset{\sim}{u})_e]. \tag{3.78}$$

A better estimate of the true image is obtained by inverse Fourier transforming the product of $F(\underset{\sim}{u})_e$ and the diffraction-limited transfer function of the telescope [Baba et al 1984 ϕ3], giving

$$s_{KT}(\underset{\sim}{x}) = \mathbf{F}^{-1}[F(\underset{\sim}{u})_e H_D(\underset{\sim}{u})]. \tag{3.79}$$

Since phase$[F(\underset{\sim}{u})]_e$ is found from a sum of phase differences, accumulated errors tend to reduce the fidelity of phase$[F(\underset{\sim}{u})]_e$ for large $|\underset{\sim}{u}|$. The choice of the magnitude of the frequency increment $\Delta\underset{\sim}{u}$ is a compromise between increased errors from summing increased numbers of noisy phase differences if $|\Delta\underset{\sim}{u}|$ is too small, and increased errors from the diminished magnitude of the Knox-Thompson transfer function (3.68) in the presence of additive noise to (3.67) if $|\Delta\underset{\sim}{u}|$ is too large. Suitable values for $|\Delta\underset{\sim}{u}|$ lie in the range of $\frac{1}{4}$ to 1 times the seeing limit $r_o/\bar{\lambda}\ell$ [cf. Knox and Thompson 1974, Bates 1982b ϕ8.4]. Fried [1979 ϕ5] suggests from theoretical considerations that $|\Delta\underset{\sim}{u}|$ should be less than about $0.4\, r_o/\bar{\lambda}\ell$. To reconstruct an object of extent E (in the image plane of the telescope, cf. ϕ2.2), the sampling theorem [Bracewell 1978 ϕ10] requires samples in Fourier space to be separated by at most 1/E. Choosing $|\Delta\underset{\sim}{u}| = \frac{1}{4}\, r_o/\bar{\lambda}\ell$ allows objects up to four times the size of the seeing disc to be imaged.

In two dimensions the existence of multiple paths of summation to each point in Fourier space allows the phase error to be reduced by appropriate choice and combination of paths. Baba et al [1984] find

that summing along radial paths (3.75, 3.76) produces better image reconstructions than summing along Cartesian paths (3.73, 3.74) because the radial paths to each $\underset{\sim}{u}$ are shorter in general. Knox [1976] averages the phases from two different paths to each $\underset{\sim}{u}$. Matrix methods which effectively combine many paths of summation to produce a least-squared-error phase estimate have been devised [Hudgin 1977, Frost et al 1979, Hunt 1979, cf. Dainty 1984 ϕ7.4.7].

The Knox-Thompson method has been pursued vigorously by Nisenson and his colleagues. Nisenson and Papaliolios [1983] [cf. Deron and Fontanella 1984] analyse the performance of the method under photon-limited conditions, following Goodman and Belsher's [1976] treatment of speckle interferometry (ϕ3.3.3). Nisenson and Papaliolios show the presence of magnitude and phase photon bias terms in the $F(\underset{\sim}{u})_e$ obtained from photon-limited speckle images, which may be predicted and removed if the number of photons per speckle image is known [cf. Papaliolios et al 1985 ϕ2]. Nisenson and others have described special-purpose hardware for Knox-Thompson processing [Nisenson and Stachnik 1975, Nisenson et al 1980, Nisenson et al 1983b, cf. Papaliolios et al 1985]. The Knox-Thompson method has been applied in astronomical practice to form images of celestial objects including the Sun's surface [Stachnik et al 1977, Noyes et al 1981, Stachnik et al 1983], planetary objects [Nisenson et al 1981, Nisenson et al 1983a, Howell and McGinn 1985], and multiple stars [Nisenson et al 1985]. The method has been used at infrared wavelengths [Chelli et al 1983, Howell and McGinn 1985].

The effects of nonisoplanatism on the Knox-Thompson method are examined experimentally by Karo and Schneiderman [1975] and theoretically by Fried [1979]. Karo and Schneiderman's [1975] simulations show severe degradation of the Knox-Thompson image under nonisoplanatic conditions, while Fried [1979] concludes that nonisoplanatism introduces spurious detail into the Knox-Thompson image. Bates and Gough [1975] note that the Knox-Thompson image is diffraction-limited only if the telescope accurately tracks the motion of the object on the celestial sphere, unlike the imaging techniques described in Sections 3.4, 3.5.3, and 3.6 which are insensitive to telescope guiding errors. O'Connor and Huang's [1981] prescription for the Knox-Thompson method initially centres the centroids of the

speckle images, which effectively autoguides the telescope and alleviates this problem. Brames and Dainty [1981] present a description of the Knox-Thompson method using complex visibility zeros (§§ 3.5.4, 6.2).

A number of variations and extensions of the Knox-Thompson method have been developed. Ehn and Nisenson [1975] mention but do not describe an improved version of the Knox-Thompson method. Aitken and Desaulniers [1979] estimate the complex spectral ratio $F(\underset{\sim}{u}+\Delta\underset{\sim}{u})\,/\,F(\underset{\sim}{u})$ from the quantity $\langle S_m^*(\underset{\sim}{u})\,S_m(\underset{\sim}{u}+\Delta\underset{\sim}{u})\rangle_M \,/\, \langle|S_m(\underset{\sim}{u})|^2\rangle_M$ formed from the ensemble of speckle images, and then multiply the spectral ratios through Fourier space in a manner similar to the addition of phase differences by the Knox-Thompson method (3.71 - 3.76). Aitken and Desaulniers claim that their method does not require observations of a reference object whereas the Knox-Thompson method, which uses speckle interferometry to estimate $|F(\underset{\sim}{u})|$, does. However their method uses approximations equivalent to those which allow speckle interferometry to estimate $|F(\underset{\sim}{u})|$ without a reference object (§3.3.2). Bruck and Sodin [1984] explore the possibility of reconstructing $F(\underset{\sim}{u})$ from $\text{phase}[F(\underset{\sim}{u})]$ alone. This is a Fourier amplitude problem [Oppenheim et al 1982, Hayes 1984; cf. Oppenheim and Lim 1981], the complement of the Fourier phase problem (§3.4.2). Bruck and Sodin's [1984] study is motivated by the possibility that $\text{phase}[F(\underset{\sim}{u})]$ may be estimated more accurately by the Knox-Thompson method than $|F(\underset{\sim}{u})|$ is by speckle interferometry. The results presented by Wirnitzer [1985 §3B] suggest that the same possibility may exist for the speckle masking method (§3.5.3). A number of composite speckle imaging schemes combining the Knox-Thompson method with other methods such as Fienup's algorithms (§3.4.2) have been described [Cocke 1980, Bates and Fright 1982, Baba et al 1984].

Aitken et al [1985] have recently described a speckle imaging method which directly estimates the phase gradient of $F(\underset{\sim}{u})$ by forming components of $\langle \nabla\text{phase}[S_m(\underset{\sim}{u})]\rangle_M$ (cf. 3.65). Components of $\nabla\text{phase}[S_m(\underset{\sim}{u})]$ for each speckle image are formed by multiplying $s_m(\underset{\sim}{x})$ by linear functions $t(\underset{\sim}{x}) = \underset{\sim}{a} + \underset{\sim}{b}.\underset{\sim}{x}$ and appealing to the derivative theorem for Fourier transforms [Bracewell 1978 pp.117,245]. These components are averaged over the ensemble of speckle images. The magnitude of $F(\underset{\sim}{u})$ is estimated by speckle interferometry (3.20).

Aitken et al [1985] present a two-dimensional computer simulation of their method and describe a proposed coherent optical processor for computing the phase gradient. Aitken et al [1986] analyse the method's signal-to-noise ratio and its behaviour with photon-limited speckle images (cf. ϕ3.3.3). They claim that the photon bias caused by photon limiting degrades the images produced by their method less than it degrades the images produced by the Knox-Thompson method. Huiser [1982a] describes a separate speckle imaging method in which components of $\nabla\text{phase}[F(\underset{\sim}{u})]$ are estimated in an ingenious manner that is similar in some respects to that of Aitken et al [1985].

3.5.3 SPECKLE MASKING (TRIPLE CORRELATION)

Weigelt and Wirnitzer [1983] and Lohmann et al [1983] describe a speckle imaging technique they call speckle masking, in which the triple correlations, or equivalently the bispectra, of speckle images are averaged (these quantities are defined below). The bispectrum of the object is then recovered from the averaged bispectrum of the speckle images by compensating for a deterministic seeing-dependent transfer function [Lohmann et al 1983], similarly to the recovery of the object power spectrum in speckle interferometry (ϕ3.3). Unlike speckle interferometry, speckle masking permits estimation of the true image because the bispectrum of $f(\underset{\sim}{x})$ contains information about both the magnitude and the phase of $F(\underset{\sim}{u})$ [Bartelt et al 1984]. Lohmann et al [1983] present prescriptions for speckle masking in both the image domain [cf. Weigelt 1977] and the Fourier domain. Speckle masking is here classified as a Fourier domain method because the more general prescription of Lohmann et al [1983] and the proof of recoverability of $F(\underset{\sim}{u})$ by Bartelt et al [1984] are both set in the Fourier domain.

Triple correlation is a third-order extension of the more familiar second-order operations of autocorrelation and cross-correlation. Theory and general applications of triple correlation are reviewed by Lohmann and Wirnitzer [1984], and many of its properties relevant to the speckle masking method are derived by Lohmann et al [1983 Appendices]. Here a brief introduction to triple correlation and the bispectrum is given before describing the speckle masking method.

Let $f_1(\underset{\sim}{x})$, $f_2(\underset{\sim}{x})$, and $f_3(\underset{\sim}{x})$ be three complex-valued functions of a K-dimensional position vector $\underset{\sim}{x}$. Their cross triple correlation $f^3_{321}(\underset{\sim}{x},\underset{\sim}{x}')$ is a complex function of a 2K-dimensional position vector $(\underset{\sim}{x},\underset{\sim}{x}')$, defined by

$$f^3_{321}(\underset{\sim}{x},\underset{\sim}{x}') = \left(f_1(\underset{\sim}{x})\ f_2(\underset{\sim}{x}+\underset{\sim}{x}')\right) \otimes f_3(\underset{\sim}{x})$$

$$= \int f_1^*(\underset{\sim}{x}'')\, f_2^*(\underset{\sim}{x}''+\underset{\sim}{x}')\, f_3(\underset{\sim}{x}+\underset{\sim}{x}'')\ .d\underset{\sim}{x}'' \ . \qquad (3.80)$$

Let $F_1(\underset{\sim}{u})$, $F_2(\underset{\sim}{u})$, and $F_3(\underset{\sim}{u})$ be the K-dimensional Fourier transforms of $f_1(\underset{\sim}{x})$, $f_2(\underset{\sim}{x})$, and $f_3(\underset{\sim}{x})$ respectively, i.e.

$$F_1(\underset{\sim}{u}) = \mathbf{F}[f_1(\underset{\sim}{x})],$$

$$F_2(\underset{\sim}{u}) = \mathbf{F}[f_2(\underset{\sim}{x})],$$

$$F_3(\underset{\sim}{u}) = \mathbf{F}[f_3(\underset{\sim}{x})]. \qquad (3.81)$$

The 2K-dimensional Fourier transform of the cross triple correlation (3.80) is called the cross bispectrum, $F^3_{321}(\underset{\sim}{u},\underset{\sim}{u}')$. It is a complex function of the 2K-dimensional position vector $(\underset{\sim}{u},\underset{\sim}{u}')$ and is equal to

$$F^3_{321}(\underset{\sim}{u},\underset{\sim}{u}') = \mathbf{F}[f^3_{321}(\underset{\sim}{x},\underset{\sim}{x}')]$$

$$= \iint f^3_{321}(\underset{\sim}{x},\underset{\sim}{x}')\ e^{-j2\pi(\underset{\sim}{u}.\underset{\sim}{x}\ +\ \underset{\sim}{u}'.\underset{\sim}{x}')}\ .d\underset{\sim}{x}.d\underset{\sim}{x}' \qquad (3.82)$$

$$= F_3(\underset{\sim}{u})\ F_2^*(-\underset{\sim}{u}')\ F_1^*(\underset{\sim}{u}+\underset{\sim}{u}'). \qquad (3.83)$$

Equation (3.83) follows directly from inserting (3.80) into (3.82) and using (3.81) and (1.12a). The (auto) triple correlation of a K-dimensional function $f(\underset{\sim}{x})$ is the 2K-dimensional function $f^3(\underset{\sim}{x},\underset{\sim}{x}')$ obtained by setting $f_1(\underset{\sim}{x})$, $f_2(\underset{\sim}{x})$, and $f_3(\underset{\sim}{x})$ equal to $f(\underset{\sim}{x})$ in (3.80), giving

$$f^3(\underset{\sim}{x},\underset{\sim}{x}') = \left(f(\underset{\sim}{x})\ f(\underset{\sim}{x}+\underset{\sim}{x}')\right) \otimes f(\underset{\sim}{x}). \qquad (3.84)$$

The bispectrum of $f(\underset{\sim}{x})$ is the 2K-dimensional Fourier transform of the auto triple correlation of $f(\underset{\sim}{x})$,

$$F^3(\underset{\sim}{u},\underset{\sim}{u}') = \mathbf{F}[f^3(\underset{\sim}{x},\underset{\sim}{x}')] = F(\underset{\sim}{u})\ F^*(-\underset{\sim}{u}')\ F^*(\underset{\sim}{u}+\underset{\sim}{u}') \qquad (3.85)$$

where $f(\underset{\sim}{x}) \leftrightarrow F(\underset{\sim}{u})$. If $f(\underset{\sim}{x})$ is real-valued, then $F(\underset{\sim}{u})$ is Hermitian,

i.e.

$$F(\underset{\sim}{u}) = F^*(-\underset{\sim}{u}), \tag{3.86}$$

and (3.85) becomes

$$F^3(\underset{\sim}{u},\underset{\sim}{u}') = F(\underset{\sim}{u})\, F(\underset{\sim}{u}')\, F(-\underset{\sim}{u}-\underset{\sim}{u}'). \tag{3.87}$$

This is the expression for a real-valued $f(\underset{\sim}{x})$ used by Lohmann, Weigelt, and others throughout the references quoted in this subsection [cf. Lohmann et al 1983 Appendix A].

Figure 3.7 illustrates the triple correlation and the bispectrum of a one-dimensional function $f(x)$. The triple correlation and the bispectrum are two-dimensional. For a two-dimensional $f(\underset{\sim}{x})$ the triple correlation and bispectrum are four-dimensional functions. The triple correlation (Figure 3.7c) and the bispectrum (Figure 3.7d) of $f(\underset{\sim}{x})$ are invariant under translations of $f(\underset{\sim}{x})$ in image space [cf. Bartelt and Wirnitzer 1985], a property they share with the autocorrelation $\mathbf{A}[f(\underset{\sim}{x})]$ and the power spectrum $|F(\underset{\sim}{u})|^2$. For a real-valued $f(\underset{\sim}{x})$ the bispectrum possesses the symmetries (easily verifiable from (3.87))

$$\begin{aligned} F^3(\underset{\sim}{u},\underset{\sim}{u}') &= F^3(\underset{\sim}{u}',\underset{\sim}{u}) \\ &= F^3(-\underset{\sim}{u}-\underset{\sim}{u}',\underset{\sim}{u}') \\ &= F^{3*}(-\underset{\sim}{u},-\underset{\sim}{u}') \end{aligned} \tag{3.88}$$

[cf. Lohmann et al 1983 Appendix A, Bartelt et al 1984 §IV]. These three equalities produce an eight-fold symmetry in the bispectrum of a real-valued $f(\underset{\sim}{x})$.

The triple correlation of a speckle image $s_m(\underset{\sim}{x})$ (3.3e,f) is denoted by $s_m^3(\underset{\sim}{x},\underset{\sim}{x}')$. The corresponding bispectrum is denoted by $S_m^3(\underset{\sim}{u},\underset{\sim}{u}')$, and is here called the speckle bispectrum. For isoplanatic speckle images (3.3e,f),

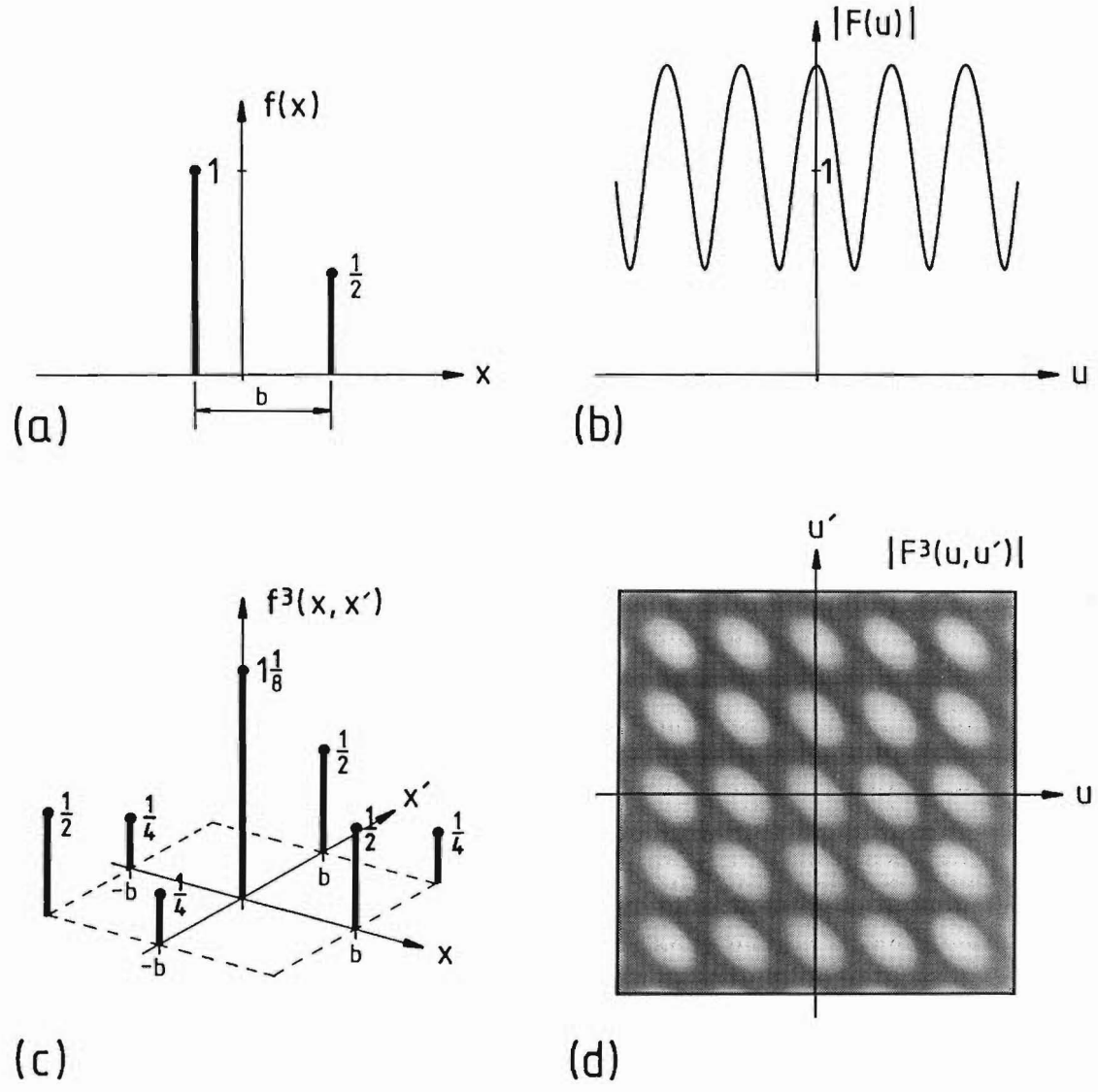


Figure 3.7: Example of the triple correlation and bispectrum of a one-dimensional function [cf. Wirnitzer 1985 Fig.1].
(a) A one-dimensional function f(x),
(b) the magnitude of the visibility of f(x),
(c) the triple correlation of f(x),
(d) the magnitude of the bispectrum of f(x) (zero magnitude is displayed as black).

$$S_m(\underset{\sim}{u},\underset{\sim}{u}') = S_m(\underset{\sim}{u})\, S_m(\underset{\sim}{u}')\, S_m(-\underset{\sim}{u}-\underset{\sim}{u}')$$

$$= F(\underset{\sim}{u})\, F(\underset{\sim}{u}')\, F(-\underset{\sim}{u}-\underset{\sim}{u}')\, H_m(\underset{\sim}{u})\, H_m(\underset{\sim}{u}')\, H_m(-\underset{\sim}{u}-\underset{\sim}{u}')$$

$$= F^3(\underset{\sim}{u},\underset{\sim}{u}')\, H_m^3(\underset{\sim}{u},\underset{\sim}{u}'), \qquad (3.89)$$

i.e. $S_m^3(\underset{\sim}{u},\underset{\sim}{u}')$ is the product of the bispectrum of the object and the bispectrum of the speckle point spread function. The product of the bispectra in (3.89) is equivalent to a 2K-dimensional convolution of triple correlations in image space,

$$s_m^3(\underset{\sim}{x},\underset{\sim}{x}') = f^3(\underset{\sim}{x},\underset{\sim}{x}') \circledcirc h_m^3(\underset{\sim}{x},\underset{\sim}{x}'), \qquad (3.90)$$

where $h_m^3(\underset{\sim}{x},\underset{\sim}{x}')$ is the triple correlation of the speckle psf $h_m(\underset{\sim}{x})$. In the speckle masking method either $S_m^3(\underset{\sim}{u},\underset{\sim}{u}')$ or $s_m^3(\underset{\sim}{x},\underset{\sim}{x}')$ are averaged over an ensemble of speckle images, yielding

$$\langle S_m^3(\underset{\sim}{u},\underset{\sim}{u}')\rangle_M = F^3(\underset{\sim}{u},\underset{\sim}{u}')\, \langle H_m^3(\underset{\sim}{u},\underset{\sim}{u}')\rangle_M , \qquad (3.91)$$

$$\langle s_m^3(\underset{\sim}{x},\underset{\sim}{x}')\rangle_M = f^3(\underset{\sim}{x},\underset{\sim}{x}') \circledcirc \langle h_m^3(\underset{\sim}{x},\underset{\sim}{x}')\rangle_M . \qquad (3.92)$$

The quantity $\langle H_m^3(\underset{\sim}{u},\underset{\sim}{u}')\rangle$ is called the speckle masking transfer function [Lohmann et al 1983 §II], and $\langle h_m^3(\underset{\sim}{x},\underset{\sim}{x}')\rangle$ is here called the speckle masking point spread function.

Lohmann et al [1983] present two prescriptions for speckle imaging via speckle masking. The first prescription operates in image space [Lohmann et al 1983 §III.A, Weigelt and Wirnitzer 1983], and is a close relative of Weigelt's [1977] earlier speckle imaging technique from whence the name "speckle masking" came to be applied to triple correlations. The second prescription operates with bispectra in Fourier space [Lohmann et al 1983 §III.B, Bartelt et al 1984]. These two approaches are described below in turn.

In image space, speckle masking is applied by finding a vector constant $\underset{\sim}{b}$, the masking vector, such that

$$f(\underset{\sim}{x})\, f(\underset{\sim}{x}+\underset{\sim}{b}) = a\,\delta(\underset{\sim}{x}) \qquad (3.93)$$

[cf. Lohmann et al 1983 Fig.2]. For a double star, for example

[Weigelt 1977], $\underset{\sim}{b}$ is the separation between the two stars (cf. Figure 3.7a). Masking vectors satisfying (3.93) exist only for suitable forms of object and are found by inspecting the autocorrelation estimate $\mathbf{A}[f(\underset{\sim}{x})]_e$ generated by speckle interferometry (3.21, 3.22). For any $\underset{\sim}{b}$ satisfying (3.93),

$$\begin{aligned} f^3(\underset{\sim}{x},\underset{\sim}{b}) &\approx a\,\delta(\underset{\sim}{x}) \otimes f(\underset{\sim}{x}) \qquad \text{from (3.84) and (3.93)} \\ &= a\,f(\underset{\sim}{x}) \end{aligned} \tag{3.94}$$

(cf. Figure 3.7c, in which scaled replicas of f(x) can be seen along the lines $u = \pm b$ and $u' = \pm b$). Computing the averaged triple correlation (3.92) with $\underset{\sim}{x}' = \underset{\sim}{b}$ yields

$$\begin{aligned} \langle s_m^3(\underset{\sim}{x},\underset{\sim}{b})\rangle_M &= f^3(\underset{\sim}{x},\underset{\sim}{b}) \odot \langle h_m^3(\underset{\sim}{x},\underset{\sim}{b})\rangle_M \\ &\approx a\,f(\underset{\sim}{x}) \odot \langle h_m^3(\underset{\sim}{x},\underset{\sim}{b})\rangle_M \,. \end{aligned} \tag{3.95}$$

Note that this is only a K-dimensional computation. An estimate of the true image of $f(\underset{\sim}{x})$ is formed by compensating for the speckle masking psf in (3.95). Lohmann et al [1983] suggest two ways of performing the compensation. The first way is to repeat the speckle masking with an unresolved object, $f_o(\underset{\sim}{x}) \approx \delta(\underset{\sim}{x})$ say, producing

$$\begin{aligned} \langle s_{om}^3(\underset{\sim}{x},\underset{\sim}{b})\rangle_{Mo} &= f^3(\underset{\sim}{x},\underset{\sim}{b}) \odot \langle h_{om}^3(\underset{\sim}{x},\underset{\sim}{b})\rangle_{Mo} \\ &\approx \delta(\underset{\sim}{x}) \odot \langle h_{om}^3(\underset{\sim}{x},\underset{\sim}{b})\rangle_{Mo} \\ &= \langle h_{om}^3(\underset{\sim}{x},\underset{\sim}{b})\rangle_{Mo} \\ &\approx \langle h_m^3(\underset{\sim}{x},\underset{\sim}{b})\rangle_M \,, \end{aligned} \tag{3.96}$$

and then to deconvolve (3.96) from (3.95). This is analogous to the standard compensation technique in speckle interferometry (ϕ3.3.1). The second way is to appropriately subtract averaged cross triple correlations of statistically independent speckle images from $\langle s_m^3(\underset{\sim}{x},\underset{\sim}{b})\rangle_M$, following the result

$$\begin{aligned} &\langle s_m^3(\underset{\sim}{x},\underset{\sim}{x}')\rangle - \langle s_{mmn}^3(\underset{\sim}{x},\underset{\sim}{x}')\rangle - \langle s_{mnm}^3(\underset{\sim}{x},\underset{\sim}{x}')\rangle \\ &\quad - \langle s_{nmm}^3(\underset{\sim}{x},\underset{\sim}{x}')\rangle + 2\langle s_{mnp}^3(\underset{\sim}{x},\underset{\sim}{x}')\rangle \\ &\quad \approx f^3(\underset{\sim}{x},\underset{\sim}{x}') \odot h_D^3(\underset{\sim}{x},\underset{\sim}{x}') \end{aligned} \tag{3.97}$$

derived by Lohmann et al [1983 Appendix B] under the assumptions of complex Gaussian statistics of the aperture plane field (ϕ2.3.1) and of very poor seeing, i.e.

$$\langle H_m(\underset{\sim}{u})\rangle \approx \delta(\underset{\sim}{u}). \tag{3.98}$$

$h_D^3(\underset{\sim}{x},\underset{\sim}{x}')$ is a diffraction-limited 2K-dimensional psf which is independent of atmospheric or telescope aberrations. Equation (3.97) is evaluated for $\underset{\sim}{x}' = \underset{\sim}{b}$ to yield a diffraction-limited estimate of $f(\underset{\sim}{x})$. Lohmann et al [1983 ϕIII.A] note that in practice it is necessary to multiply the first term of (3.97) by a scaling constant $C \approx 1$ to obtain a good reconstruction of $f(\underset{\sim}{x})$ under non-Gaussian statistics of the aperture plane field. I believe that the procedure described by (3.97) is exactly analogous to Welter and Worden's [1978] method of compensating for the Labeyrie transfer function (ϕ3.3.2, (3.25)) and suffers a similar attenuation of spatial frequencies of $f(\underset{\sim}{x})$ below the seeing limit. The need to apply a correction factor $C \approx 1$ is then analogous to the need for a similar correction factor when using long exposure subtraction in speckle interferometry (3.28, 3.29). Note that the assumption (3.98) used to derive (3.97) implies that the seeing limit is zero and hence no spatial frequencies are attenuated in (3.97).

The Fourier space implementation of speckle masking [Lohmann et al 1983 ϕIII.B, Bartelt et al 1984] separately estimates the magnitude and the phase of $F(\underset{\sim}{u})$ from the averaged speckle bispectrum (3.91). The value of (3.91) along any of the lines $\underset{\sim}{u} = \underset{\sim}{0}$, $\underset{\sim}{u}' = \underset{\sim}{0}$, or $\underset{\sim}{u} + \underset{\sim}{u}' = \underset{\sim}{0}$ allows $\langle|S_m(\underset{\sim}{u})|^2\rangle_M$ to be determined directly since, taking $\underset{\sim}{u} = \underset{\sim}{0}$ for example and noting that $s_m(\underset{\sim}{x})$ is real,

$$\begin{aligned}\langle S_m^3(\underset{\sim}{u},\underset{\sim}{0})\rangle_M &= \langle S_m(\underset{\sim}{u})\, S_m(\underset{\sim}{0})\, S_m(-\underset{\sim}{u})\rangle_M \\ &= S_m(\underset{\sim}{0})\, \langle S_m(\underset{\sim}{u})\, S_m^*(\underset{\sim}{u})\rangle_M \\ &= (\text{real constant}).\langle|S_m(\underset{\sim}{u})|^2\rangle_M\end{aligned} \tag{3.99}$$

$|F(\underset{\sim}{u})|$ can then be estimated from (3.99) as in speckle interferometry (ϕ3.3). Bartelt et al [1984 ϕIV.B] show that it is also possible to estimate $|F(\underset{\sim}{u})|$ from parts of the bispectrum other than the lines $\underset{\sim}{u} = \underset{\sim}{0}$, $\underset{\sim}{u}' = \underset{\sim}{0}$, $\underset{\sim}{u} + \underset{\sim}{u}' = \underset{\sim}{0}$, which allows an improved estimate of $|F(\underset{\sim}{u})|$

to be obtained if the bispectrum itself is contaminated by noise. The phase of $F(\underset{\sim}{u})$ can be obtained from (3.91) without any compensation for the seeing since the speckle masking transfer function is real-valued [Lohmann et al 1983 Appendix B], i.e.

$$\begin{aligned}\text{phase}[\langle S_m^3(\underset{\sim}{u},\underset{\sim}{u}')\rangle_M] &= \text{phase}[F^3(\underset{\sim}{u},\underset{\sim}{u}')] + \text{phase}[\langle H_m^3(\underset{\sim}{u},\underset{\sim}{u}')\rangle_M] \\ &\approx \text{phase}[F^3(\underset{\sim}{u},\underset{\sim}{u}')]. \end{aligned} \tag{3.100}$$

A recursive algorithm for determining $\text{phase}[F(\underset{\sim}{u})]$ from $\text{phase}[F^3(\underset{\sim}{u},\underset{\sim}{u}')]$ is described in the following paragraph. (3.100) allows $\langle S_m^3(\underset{\sim}{u},\underset{\sim}{u}')\rangle_M$ to be used in place of $F^3(\underset{\sim}{u},\underset{\sim}{u}')$ in the algorithm. This produces an estimate of $\text{phase}[F(\underset{\sim}{u})]$ which is combined with the estimated $|F(\underset{\sim}{u})|$ and inverse Fourier transformed to yield an estimate of the true image.

Bartelt et al [1984 §IV.A] and Lohmann et al [1983 §III.B] describe recursive algorithms for extracting $\text{phase}[F(\underset{\sim}{u})]$ from the bispectrum of $f(\underset{\sim}{x})$. For a real-valued $f(\underset{\sim}{x})$, one has

$$F^3(\underset{\sim}{u},\underset{\sim}{u}') = F(\underset{\sim}{u})\,F(\underset{\sim}{u}')\,F(-\underset{\sim}{u}-\underset{\sim}{u}') = F(\underset{\sim}{u})\,F(\underset{\sim}{u}')\,F^*(\underset{\sim}{u}+\underset{\sim}{u}') \tag{3.101}$$

and therefore

$$\begin{aligned}\text{phase}[F^3(\underset{\sim}{u},\underset{\sim}{u}')] = \text{phase}[F(\underset{\sim}{u})] &+ \text{phase}[F(-\underset{\sim}{u}')] \\ &- \text{phase}[F(\underset{\sim}{u}+\underset{\sim}{u}')]. \end{aligned} \tag{3.102}$$

To introduce a compact notation, let

$$\phi_p = \text{phase}[F(p\Delta\underset{\sim}{u})] \tag{3.103a}$$

and

$$\phi_{p;q} = \text{phase}[F^3(p\Delta\underset{\sim}{u},q\Delta\underset{\sim}{u}')], \tag{3.103b}$$

where p and q are integers and $\Delta\underset{\sim}{u}$ and $\Delta\underset{\sim}{u}'$ are suitable sampling intervals in Fourier space. The $\phi_{p;q}$ are given data and the ϕ_p are the phases to be determined from the given data. Note that

$$\phi_{-p} = -\phi_p \tag{3.104}$$

(cf. 3.86) since $f(\underset{\sim}{x})$ is real. Using this notation, (3.102) becomes

$$\phi_{p;q} = \phi_p + \phi_q - \phi_{p+q} \Rightarrow \phi_{p+q} = \phi_p + \phi_q - \phi_{p;q} . \tag{3.105}$$

Equation (3.105) can be used recursively to solve for ϕ_p, for any fixed value of $q \neq 0$. Let $p = nq$ for integer n, with q equal to some fixed value. Equation (3.105) becomes

$$\phi_{(n+1)q} = \phi_{nq} + \phi_q - \phi_{nq;q} . \tag{3.106}$$

Since $f(\underset{\sim}{x})$ is real,

$$\phi_0 = 0 . \tag{3.107}$$

Setting $n = 0$ in (3.106) gives

$$\phi_q = 0 + \phi_q - \phi_{0;q} . \tag{3.108a}$$

$\phi_{0;q}$ is zero because $F^3(\underset{\sim}{0},\underset{\sim}{u}')$ is real-valued for all $\underset{\sim}{u}'$ and real-valued $f(\underset{\sim}{x})$ (3.99). Hence ϕ_q is indeterminate from (3.108a) and must be chosen arbitrarily (a suitable choice is $\phi_q = 0$). Deferring the choice of value for ϕ_q for the moment and setting $n = 1,2,\ldots$ in (3.106) gives

$$\phi_{2q} = \phi_q + \phi_q - \phi_{q;q} , \tag{3.108b}$$

$$\phi_{3q} = \phi_{2q} + \phi_q - \phi_{2q;q}$$

$$= \phi_q + \phi_q + \phi_q - \phi_{2q;q} - \phi_{q;q} , \tag{3.108c}$$

$$\ldots ,$$

which can be compactly expressed as the recursive formula

$$\phi_{nq} = n\,\phi_q + \sum_{m=0}^{n-1} \phi_{mq;q} \tag{3.109}$$

for $n \geq 1$. Equation (3.109) is used with (3.107) and (3.104) to determine phase$[F(\underset{\sim}{u})]$ at points in Fourier space spaced by $q\Delta\underset{\sim}{u}$. Using $q = 1$ allows all ϕ_p (3.103a) to be found from $\phi_{p;q}$, for as many values of p as the measured data permit (e.g. to the diffraction limit

of the telescope in speckle masking). Using $q = 2,3,\ldots$ allows independent estimates of some of the ϕ_p to be made from different sets of ϕ_{pq}, thereby improving the signal-to-noise ratio of the ϕ_p for noisy $\phi_{p;q}$ [Bartelt et al 1984 §V]. The arbitrary choice of ϕ_q in (3.108) produces an arbitrary linear phase factor $n\phi_q$ in ϕ_{nq} (3.109), which corresponds to an arbitrary translation of the estimated true image in image space. This is a consequence of the previously noted invariance of triple correlations and bispectra to translations of $f(\underset{\sim}{x})$.

Astronomical observations by speckle masking are reported by Weigelt et al [1985b] and Hofmann et al [198_] and also by Weigelt and Wirnitzer [1983] and Lohmann et al [1983], the object in each case being a multiple star.

The Fourier domain implementation of speckle masking, though apparently more general than the image domain implementation, has the disadvantage of doubling the dimensionality of the position vector $\underset{\sim}{u}$. Two-dimensional speckle images require the evaluation of four-dimensional bispectra, which is computationally expensive. Lohmann et al [1983] avoid four-dimensional bispectra in a double star observation by performing speckle masking on one-dimensional projections of the two-dimensional speckle images. This yields estimates of the corresponding projections of the true image, which allow a good estimate of the true image to be formed if sufficiently many projections are available [cf. Lewitt and Bates 1978, Lewitt et al 1978]. Hofmann and Weigelt [198_] propose one-dimensional speckle masking with a space-borne interferometer having an essentially one-dimensional aperture. Special-purpose hardware for performing the triple correlation is reviewed by Lohmann and Wirnitzer [1984 §IV.A]; see also Cohen [1985].

Speckle masking is a third-order process [cf. Bates and Gough 1975 §IV.A] since three quantities are multiplied together in Fourier space (cf. the first line of (3.89)). The signal-to-noise ratio in each speckle bispectrum is therefore very poor if the signal-to-noise ratio of the speckle images, and hence of their speckle visibilities, is poor. The averaging of redundant information through the bispectrum when determining $|F(\underset{\sim}{u})|$ and $\text{phase}[F(\underset{\sim}{u})]$ can alleviate the

poor signal-to-noise ratio in the averaged speckle bispectrum [cf. Bartelt et al 1984 φV, Wirnitzer 1985 φ3C]. The behaviour of the bispectrum under photon-limited conditions is examined theoretically by Wirnitzer [1985], who follows Goodman and Belsher's [1976] treatment of speckle interferometry (φ3.3.3) and shows the existence of photon bias terms in the averaged speckle bispectrum which can be estimated and removed by subtraction. Wirnitzer [1985] concludes that speckle masking can usefully estimate the true image whenever speckle interferometry can usefully estimate the autocorrelation of the object.

3.5.4 COMPLEX ZERO METHODS

In this subsection, Section 6.2 is briefly anticipated in order to describe speckle imaging methods which invoke the concept of complex zeros of the visibility of a function.

The visibility $F(\underset{\sim}{u})$ of a function $f(\underset{\sim}{x})$ (1.12a) can be analytically continued into the complex domain $\underset{\sim}{w} = \underset{\sim}{u} + j\underset{\sim}{v}$ by replacing $\underset{\sim}{u}$ with $\underset{\sim}{w}$ in (1.12a), giving

$$F(\underset{\sim}{w}) = \int f(\underset{\sim}{x})\, e^{-j2\pi \underset{\sim}{w}.\underset{\sim}{x}} .d\underset{\sim}{x} . \qquad (3.110)$$

If $f(\underset{\sim}{x})$ is of finite extent then $F(\underset{\sim}{w})$ is characterised up to a complex scaling constant by its complex zeros, which are the values $\{\underset{\sim}{\omega}\}$ of $\underset{\sim}{w}$ for which $F(\underset{\sim}{w}) = 0$. The complex zeros of a one-dimensional $F(\underset{\sim}{w})$ are a set of isolated points on the complex plane, of which only a finite number are required to describe $f(\underset{\sim}{x})$ to any desired accuracy. The complex zeros of the visibility of a function are an alternative to the more familiar representation of the function by samples of its value in image space or Fourier space. They serve as a source of useful alternative insights and computational methods.

Walker [1981b] [cf. Dainty 1984 φ7.4.4] proposes a speckle imaging method that employs his earlier published Fourier phase retrieval technique based on exponential apodisation in the image domain [Walker 1981a]. Here the phase retrieval technique and then its application to speckle imaging are described.

Walker's [1981a] Fourier phase retrieval technique is described using complex visibility zeros as follows. The analytic continuation into the complex domain of the power spectrum $\Phi_f(\underset{\sim}{u}) = F(\underset{\sim}{u})\ F^*(\underset{\sim}{u})$ of a function $f(\underset{\sim}{x})$ is given by

$$\Phi_f(\underset{\sim}{w}) = F(\underset{\sim}{w})\ F^*(\underset{\sim}{w}^*), \tag{3.111}$$

as can be easily verified by substituting

$$\mathbf{A}[f(\underset{\sim}{x})] = f(\underset{\sim}{x}) \odot f^*(-\underset{\sim}{x}) \leftrightarrow \Phi_f(\underset{\sim}{u}) \tag{3.112}$$

into (3.110). The complex zeros of $F(\underset{\sim}{w})$ form a set $\{\underset{\sim}{\omega}\}$ and the complex zeros of $F^*(\underset{\sim}{w}^*)$ form a second set $\{\underset{\sim}{\omega}^*\}$ in which each complex zero is the complex conjugate of the corresponding complex zero in the first set. The complex zeros of $\Phi_f(\underset{\sim}{w})$ are the union of these two sets, $\{\underset{\sim}{\omega},\underset{\sim}{\omega}^*\}$. Fourier phase retrieval implies choosing from each complex conjugate pair of complex zeros in $\{\underset{\sim}{\omega},\underset{\sim}{\omega}^*\}$ the complex zero belonging to $F(\underset{\sim}{w})$. If $f(\underset{\sim}{x})$ is multiplied by a real exponential function $e^{-j2\pi\underset{\sim}{b}.\underset{\sim}{x}}$, where $\underset{\sim}{b}$ is a vector constant, the resulting exponentially apodised function

$$f'(\underset{\sim}{x}) = f(\underset{\sim}{x})\ e^{-2\pi\underset{\sim}{b}.\underset{\sim}{x}} \tag{3.113}$$

has an analytically continued visibility given by

$$F'(\underset{\sim}{w}) = F(\underset{\sim}{w} - j\underset{\sim}{b}), \tag{3.114}$$

as is shown by substituting (3.113) into (3.110), which is a translated version of $F(\underset{\sim}{w})$. The analytically continued power spectrum of $f'(\underset{\sim}{x})$ is given by

$$\begin{aligned}\Phi'_f(\underset{\sim}{w}) &= F'(\underset{\sim}{w})\ F'^*(\underset{\sim}{w}^*) \\ &= F(\underset{\sim}{w} - j\underset{\sim}{b})\ F^*(\underset{\sim}{w}^* + j\underset{\sim}{b})\end{aligned} \tag{3.115}$$

and has the complex zeros $\{\underset{\sim}{\omega} + j\underset{\sim}{b},\underset{\sim}{\omega}^* - j\underset{\sim}{b}\}$. The complex zeros of $\Phi'_f(\underset{\sim}{w})$ belonging to $F(\underset{\sim}{w})$ can be identified by comparing $\{\underset{\sim}{\omega},\underset{\sim}{\omega}^*\}$ from $\Phi_f(\underset{\sim}{w})$ with $\{\underset{\sim}{\omega} + j\underset{\sim}{b},\underset{\sim}{\omega}^* - j\underset{\sim}{b}\}$ from $\Phi'_f(\underset{\sim}{w})$ and choosing from each complex conjugate pair in $\{\underset{\sim}{\omega},\underset{\sim}{\omega}^*\}$ the complex zero which is translated by $+j\underset{\sim}{b}$

in $\{\underset{\sim}{\omega} + j\underset{\sim}{b}, \underset{\sim}{\omega}^* - j\underset{\sim}{b}\}$. The choice is illustrated for one-dimensional $f(\underset{\sim}{x})$ in Figure 3.8. Walker [1981a] remarks that phase retrieval by exponential apodisation can be implemented either with complex zero location algorithms and the procedure just described or by an iterative procedure [Walker 1981b] similar to Fienup's algorithms (ϕ3.4.2).

The application of Walker's [1981a] Fourier phase retrieval technique to speckle imaging is described by Walker [1981b]. Given an ensemble of speckle images, speckle interferometry (ϕ3.3) is performed on the speckle images to produce an estimate of $|F(\underset{\sim}{u})|^2 = \Phi_f(\underset{\sim}{u})$, and hence of $\Phi_f(\underset{\sim}{w})$. In addition to this, each speckle image is exponentially apodised to form an ensemble of modified speckle images

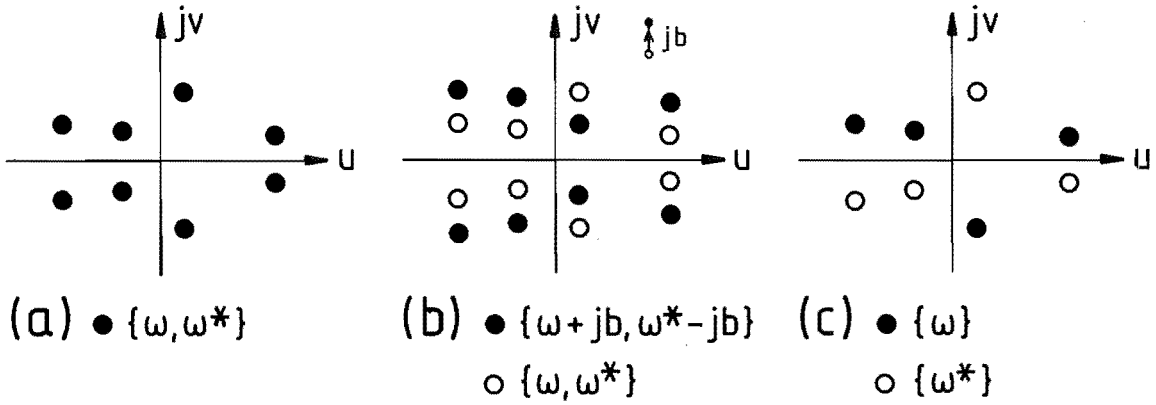


Figure 3.8: Walker's [1981a] Fourier phase retrieval technique illustrated for a one-dimensional object.

(a) Complex zeros of the power spectrum $\Phi_f(w) = F(w)F^*(w^*)$ of the object f(x) (black dots).

(b) Complex zeros of the power spectrum $\Phi_f'(w)$ of the exponentially apodised object $f'(x) = f(x)\, e^{-2\pi bx}$ (black dots).

(c) The complex zeros belonging to F(w) in (a) are the ones which move upwards (by jb) in going from (a) to (b) (black dots in (c)).

$$s'_m(\underset{\sim}{x}) = s_m(\underset{\sim}{x})\, e^{-2\pi \underset{\sim}{b}.\underset{\sim}{x}}$$

$$= \left(f(\underset{\sim}{x}) \odot h_m(\underset{\sim}{x})\right) e^{-2\pi \underset{\sim}{b}.\underset{\sim}{x}}$$

$$= \left(f(\underset{\sim}{x})\, e^{-2\pi \underset{\sim}{b}.\underset{\sim}{x}}\right) \odot \left(h_m(\underset{\sim}{x})\, e^{-2\pi \underset{\sim}{b}.\underset{\sim}{x}}\right)$$

$$= f'(\underset{\sim}{x}) \odot h'_m(\underset{\sim}{x}) \qquad (3.116a)$$

$$\leftrightarrow F'(\underset{\sim}{u})\, H'_m(\underset{\sim}{u}) \qquad (3.116b)$$

[cf. Walker 1982 Appendix]. Speckle interferometry is performed on the ensemble $\{s'_m(\underset{\sim}{x})\}$, yielding the result

$$\langle |S'_m(\underset{\sim}{u})|^2 \rangle_M = |F'(\underset{\sim}{u})|^2 \langle |H'_m(\underset{\sim}{u})|^2 \rangle_M \qquad (3.117)$$

in which a modified transfer function $\langle |H'_m(\underset{\sim}{u})|^2 \rangle_M$ appears. $\langle |H'_m(\underset{\sim}{u})|^2 \rangle_M$ can be compensated for in the same way that $\langle |H_m(\underset{\sim}{u})|^2 \rangle_M$ is with $\{s_m(\underset{\sim}{x})\}$, for example by observing an unresolved reference object. $|F'(\underset{\sim}{u})|^2$ and hence $\Phi_f(\underset{\sim}{w})$ are then estimated from (3.117) and the phase retrieval algorithm described above is applied. Walker [1982] presents computer simulated examples of this procedure for an extended two-dimensional object at high and low light levels.

The zero-and-add speckle imaging method recently proposed by Bates et al [1985] [cf. Davey et al 198_, Sinton et al 198_] computes the complex zeros $\{\omega\}_{sm}$ of the visibilities of individual one-dimensional speckle images $s_m(x)$. Since

$$S_m(w) = F(w) H_m(w) \qquad (3.118)$$

under perfectly isoplanatic conditions (cf. 3.3f), the complex zeros of $S_m(u)$ are the union of the complex zeros $\{\omega\}_f$ of $F(u)$ and the complex zeros $\{\omega\}_{hm}$ of $H_m(u)$, i.e.

$$\{\omega\}_{sm} = \{\omega\}_f \cup \{\omega\}_{hm} . \qquad (3.119)$$

The set $\{\omega\}_f$ is identical for all speckle images, whereas the members of $\{\omega\}_{hm}$ randomly change position on the complex plane from speckle image to speckle image because $H_m(u)$ is random. Comparing the $\{\omega\}_{sm}$

from different speckle images allows $\{\omega\}_f$ to be identified and hence an image of f(x) to be reconstructed. Deviations from perfect isoplanatism in the speckle images, expressible as an additive contamination $C_m(w)$ in

$$S_m(w) = F(w)H_m(w) + C_m(w), \tag{3.120}$$

cause the complex zeros in $\{\omega\}_{sm}$ associated with $\{\omega\}_f$ to change position randomly from speckle image to speckle image. They remain identifiable if their movements are less random than the movements of the complex zeros associated with $\{\omega\}_{hm}$. This is the case provided the contamination is not too severe. Zero-and-add is described in detail in Chapter 6 of this thesis, including consideration of the effects of various forms of contamination. Sinton et al [198_] apply zero-and-add to computer simulated speckle imaging in two dimensions. The need to compute complex zeros for each speckle image makes zero-and-add in its present form computationally expensive. However it holds the promise of being able to accurately recover F(u) under isoplanatic conditions without having to measure or estimate any seeing-dependent transfer function.

A final note for this subsection is that Brames and Dainty [1981] provide an interpretation of the object cross-spectrum $F^*(\underset{\sim}{u})\ F(\underset{\sim}{u} + \Delta\underset{\sim}{u})$ estimated by the Knox-Thompson method (ϕ3.5.2) in terms of complex visibility zeros. The analytically continued cross spectrum $F^*(\underset{\sim}{w}^*)\ F(\underset{\sim}{w} + \Delta\underset{\sim}{u})$ has complex zeros $\{\underset{\sim}{\omega}^*, \underset{\sim}{\omega} - \Delta\underset{\sim}{u}\}$. Comparing these complex zeros with the complex zeros $\{\underset{\sim}{\omega}, \underset{\sim}{\omega}^*\}$ of the power spectrum $\Phi_f(\underset{\sim}{w})$ allows the complex zeros $\{\underset{\sim}{\omega}\}$ of the object to be identified as the ones which differ in position by $\Delta\underset{\sim}{u}$ between the two sets.

3.6 SPECKLE IMAGING: IMAGE DOMAIN METHODS

A number of speckle imaging techniques have been proposed which operate in the image domain by means of correlation or shift-and-add operations on speckle images. Such techniques have acquired the generic title of shift-and-add techniques [e.g. Dainty 1984 ϕ7.4.5], though the shift-and-add method (ϕ3.6.2) of Bates and Cady

[1980] was not the first such technique to be proposed (cf. ϕ3.6.1). The common feature of all shift-and-add techniques is the location of individual bright speckles in the speckle images and the shifting and adding together of the speckle images with the chosen bright speckles at the centre of image space. Mertz [1984 ϕV] provides a conceptual link with the Fourier domain by noting that shifting and adding bright speckles together amounts to correcting, at high spatial frequencies only, the atmospheric phase perturbations of the object's visibility.

3.6.1 LYNDS-WORDEN-HARVEY METHOD

Lynds et al [1976] consider an approximation to the speckle point spread function consisting of the diffraction-limited psf of the telescope convolved with a collection of delta functions of various positions and strengths,

$$d_m(\underset{\sim}{x}) = \sum_{i=1}^{I_m} a'_{mi}\, \delta(\underset{\sim}{x} - \underset{\sim}{x}'_{mi}), \tag{3.121}$$

representing the positions and brightnesses of the individual speckles in $h_m(\underset{\sim}{x})$. The speckle psf is expressed as

$$h_m(\underset{\sim}{x}) = h_D(\underset{\sim}{x}) \circledcirc d_m(\underset{\sim}{x}) + n_m(\underset{\sim}{x}), \tag{3.122}$$

where $h_D(\underset{\sim}{x}) \circledcirc d_m(\underset{\sim}{x})$ is the above-mentioned approximation to $h_m(\underset{\sim}{x})$ and $n_m(\underset{\sim}{x})$ is a residual noise term comprising the differences in form between $h_D(\underset{\sim}{x})$ and each speckle in $h_m(\underset{\sim}{x})$ plus any dim speckles of $h_m(\underset{\sim}{x})$ not included in $d_m(\underset{\sim}{x})$. The speckle image is then represented by

$$\begin{aligned} s_m(\underset{\sim}{x}) &= f(\underset{\sim}{x}) \circledcirc \left[h_D(\underset{\sim}{x}) \circledcirc d_m(\underset{\sim}{x}) + n_m(\underset{\sim}{x})\right] \\ &= f(\underset{\sim}{x}) \circledcirc h_D(\underset{\sim}{x}) \circledcirc d_m(\underset{\sim}{x}) + \tilde{n}_m(\underset{\sim}{x}) \end{aligned} \tag{3.123}$$

where $\tilde{n}_m(\underset{\sim}{x})$ is a modified noise term. Equation (3.123) expresses the intuitive interpretation of a speckle image as the superposition of a number of translated, weighted, and distorted replicas of the true image, one replica for each speckle in $h_m(\underset{\sim}{x})$. The distortions of the replicas are contained in $\tilde{n}_m(\underset{\sim}{x})$. When the angular extent of the true image is not much larger than the effective extent of $h_D(\underset{\sim}{x})$, the

speckle image retains a speckled appearance (cf. ϕ2.5.1) and the form of the true image is visually apparent as the form of the individual speckles in $s_m(\underset{\sim}{x})$. Figure 1 of Worden et al [1976] serves as a good example of this: the speckles of an unresolved star are small, the speckles of a resolved star are large, and the speckles of a double star occur in pairs.

For a particular speckle psf $h_m(\underset{\sim}{x})$, $d_m(\underset{\sim}{x})$ corresponds to the positions and brightnesses of the bright speckles in the speckle image of a point object. Lynds et al [1976] observe that these positions and brightnesses are also recoverable by inspection of the speckle image of a resolved object provided the spatial extent of the object is sufficiently small, perhaps no more than four or five times the Rayleigh limit (ϕ1.6.6) [Worden et al 1976]. For such barely resolved objects, e.g. supergiant stars relatively close to Earth, the individual speckles of $s_m(\underset{\sim}{x})$ do not merge together. This allows $h_m(\underset{\sim}{x})$ to be estimated directly from $s_m(\underset{\sim}{x})$. The estimated $h_m(\underset{\sim}{x})$ is described by an estimated $d_m(\underset{\sim}{x})$,

$$d_m(\underset{\sim}{x})_e = \sum_{i=1}^{I_m} a_{mi}\, \delta(\underset{\sim}{x} - \underset{\sim}{x}_{mi}), \tag{3.124}$$

obtained from the measured positions $\underset{\sim}{x}_{mi}$ and brightnesses a_{mi} of the bright speckles in $s_m(\underset{\sim}{x})$. Deconvolving $d_m(\underset{\sim}{x})_e$ from $s_m(\underset{\sim}{x})$, assuming $d_m(\underset{\sim}{x})_e \approx d_m(\underset{\sim}{x})$, provides an estimate of the true image,

$$\begin{aligned} s_m(\underset{\sim}{x}) \odot^{-1} d_m(\underset{\sim}{x})_e &= f(\underset{\sim}{x}) \odot h_D(\underset{\sim}{x}) \odot \left[d_m(\underset{\sim}{x}) \odot^{-1} d_m(\underset{\sim}{x})_e \right] \\ &\qquad + \tilde{n}_m(\underset{\sim}{x}) \odot^{-1} d_m(\underset{\sim}{x})_e \\ &\approx f(\underset{\sim}{x}) \odot h_D(\underset{\sim}{x}) \odot \delta(\underset{\sim}{x}) + \tilde{\tilde{n}}_m(\underset{\sim}{x}) \\ &\approx s_D(\underset{\sim}{x}) + \tilde{\tilde{n}}_m(\underset{\sim}{x}) \end{aligned} \tag{3.125}$$

where $\odot^{-1}$ represents the deconvolution operation. Lynds et al [1976 ϕIII] note that deconvolution in the Fourier domain produces "minor spurious imperfections in the background" of the resulting image. They propose an alternative approach which is to cross-correlate $d_m(\underset{\sim}{x})_e$ with $s_m(\underset{\sim}{x})$, forming the sum of all of the bright speckles of $s_m(\underset{\sim}{x})$ at the centre of image space on top of a background of cross-correlation terms. This can be considered an approximate form of

deconvolution by $d_m(\underset{\sim}{x})_e$. The cross-correlation yields

$$\begin{aligned} s_{LWHm}(\underset{\sim}{x}) &= d_m(\underset{\sim}{x})_e \otimes s_m(\underset{\sim}{x}) \\ &= f(\underset{\sim}{x}) \odot h_D(\underset{\sim}{x}) \odot \left[d_m(\underset{\sim}{x})_e \otimes d_m(\underset{\sim}{x}) \right] + d_m(\underset{\sim}{x})_e \otimes \tilde{n}_m(\underset{\sim}{x}) \\ &= s_D(\underset{\sim}{x}) \odot \Bigg(\sum_{i=1}^{I_m} a_{mi} a'_{mi} \delta(\underset{\sim}{x} + \underset{\sim}{x}_{mi} - \underset{\sim}{x}'_{mi}) \\ &\quad + \sum_{p=1}^{I_m} \sum_{\substack{q=1 \\ q \neq p}}^{I_m} a_{mp} a'_{mq} \delta(\underset{\sim}{x} + \underset{\sim}{x}_{mp} - \underset{\sim}{x}'_{mq}) \Bigg) + d_m(\underset{\sim}{x})_e \otimes \tilde{n}_m(\underset{\sim}{x}) \\ &= s_D(\underset{\sim}{x}) \odot \sum_{i=1}^{I_m} a_{mi} a'_{mi} \delta(\underset{\sim}{x} + \underset{\sim}{x}_{mi} - \underset{\sim}{x}'_{mi}) + s_D(\underset{\sim}{x}) \odot c_m(\underset{\sim}{x}) \\ &\quad + d_m(\underset{\sim}{x})_e \otimes \tilde{n}_m(\underset{\sim}{x}). \end{aligned} \tag{3.126}$$

If $d_m(\underset{\sim}{x})_e$ is a good estimate of $d_m(\underset{\sim}{x})$, in the sense that

$$\underset{\sim}{x}_{mi} \approx \underset{\sim}{x}'_{mi}, \tag{3.127}$$

then (3.126) becomes

$$\begin{aligned} s_{LWHm}(\underset{\sim}{x}) &\approx s_D(\underset{\sim}{x}) \odot \sum_{i=1}^{I_m} a_{mi} a'_{mi} \delta(\underset{\sim}{x}) + s_D(\underset{\sim}{x}) \odot c_m(\underset{\sim}{x}) \\ &\quad + d_m(\underset{\sim}{x})_e \otimes \tilde{n}_m(\underset{\sim}{x}) \\ &= A_m s_D(\underset{\sim}{x}) + \tilde{c}_m(\underset{\sim}{x}) + d_m(\underset{\sim}{x})_e \otimes \tilde{n}_m(\underset{\sim}{x}). \end{aligned} \tag{3.128}$$

$\tilde{c}_m(\underset{\sim}{x})$ is the convolution of $s_D(\underset{\sim}{x})$ with all of the cross-product terms from $d_m(\underset{\sim}{x})_e \otimes d_m(\underset{\sim}{x})$. The autoproduct terms from $d_m(\underset{\sim}{x})_e \otimes d_m(\underset{\sim}{x})$ occur near the origin whereas the cross-product terms are spread out in image space, making the magnitude of $A_m s_D(\underset{\sim}{x})$ large compared to that of $\tilde{c}_m(\underset{\sim}{x})$ within the effective extent of $s_D(\underset{\sim}{x})$. $s_{LWHm}(\underset{\sim}{x})$ is averaged over the ensemble of speckle images to yield

$$\begin{aligned} s_{LWH}(\underset{\sim}{x}) &= \langle s_{LWHm}(\underset{\sim}{x}) \rangle_M \\ &\approx \langle A_m \rangle_M s_D(\underset{\sim}{x}) + \langle \tilde{c}_m(\underset{\sim}{x}) \rangle_M + \langle d_m(\underset{\sim}{x})_e \otimes \tilde{n}_m(\underset{\sim}{x}) \rangle_M \\ &= A s_D(\underset{\sim}{x}) + \tilde{\tilde{c}}(\underset{\sim}{x}) + \tilde{\tilde{n}}(\underset{\sim}{x}), \end{aligned} \tag{3.129}$$

which has the form of the true image (provided (3.127) holds) superimposed on a "fluctuating background consisting mostly of cross-product terms from the correlation" [Lynds et al 1976 §III]. This background is analogous to the fog generated by shift-and-add (§3.6.2).

Lynds et al [1976] subject each speckle image to a prefiltering step to eliminate the fog in $s_{LWH}(\underset{\sim}{x})$. They fit a two-dimensional Gaussian function to each $s_m(\underset{\sim}{x})$ and then subtract the Gaussian from $s_m'(\underset{\sim}{x})$ to leave a modified speckle image $s_m'(\underset{\sim}{x})$ in which the low spatial frequency components are suppressed. This step is akin to long exposure subtraction in speckle interferometry (§3.3.2). They then construct $d_m'(\underset{\sim}{x})_e$ from the positions and brightnesses of all speckles in $s_m'(\underset{\sim}{x})$ brighter than about 60% of the brightest speckle in $s_m'(\underset{\sim}{x})$ [Worden et al 1976], and compute the Lynds-Worden-Harvey image

$$s'_{LWHm}(\underset{\sim}{x}) = d_m'(\underset{\sim}{x})_e \otimes s_m'(\underset{\sim}{x}) \approx A_m' s_D(\underset{\sim}{x}) \tag{3.130}$$

in which the background turns out to be almost completely suppressed [Lynds et al 1976 Fig.3, cf. Welter and Worden 1980 §III]. The (composite) Lynds-Worden-Harvey image is formed by averaging $s'_{LWHm}(\underset{\sim}{x})$ over all available speckle images, giving

$$s'_{LWH}(\underset{\sim}{x}) = \langle s'_{LWHm}(\underset{\sim}{x}) \rangle_M \approx A' s_D(\underset{\sim}{x}). \tag{3.131}$$

The similarity of $s'_{LWH}(\underset{\sim}{x})$ to the true image, as implied by (3.131), depends on the absence of fine detail in the background of (3.129) in the region of image space occupied by $s_D(\underset{\sim}{x})$. Lynds et al [1976] are optimistic that this is so since the Lynds-Worden-Harvey image of an unresolved star is qualitatively similar to the diffraction-limited psf of the telescope [cf. Welter and Worden 1980 §III]. Others have expressed disquiet on this point [McDonnell and Bates 1976, Labeyrie 1978].

The Lynds-Worden-Harvey method in its original form [Lynds et al 1976] is applicable only to objects bright enough to form speckle images without appreciable photon limiting, i.e. the average number of photons per speckle (§3.3.3) is large. Because of the averaging of many speckles per speckle image, relatively few independent speckle

images are needed to reduce the variance of $s'_{LWH}(\underset{\sim}{x})$ (3.131) to an acceptable level. Welter and Worden [1980 ∅II] use about twenty speckle images, cf. the few hundred typical for speckle interferometry (∅3.3.3). Worden [1975, 1976], Lynds et al [1976], Wilkerson and Worden [1977], and Welter and Worden [1980] estimate the radii and limb darkening of a number of supergiant stars from radially averaged profiles of Lynds-Worden-Harvey images. Two-dimensional pictures of supergiant stars, notably Betelgeuse, are presented by Lynds et al [1974] [cf. Dainty 1975b ∅7.4.2, Dainty 1984 ∅7.4.5], Lynds et al [1976], Worden et al [1976], and Welter and Worden [1980]. The map of the surface of Betelgeuse produced by Lynds et al [1974] showing apparent surface detail caused considerable excitement at the time it appeared [cf. --- 1975, Irwin 1975]. Subsequent studies of Betelgeuse by Wilkerson and Worden [1977] and Welter and Worden [1980] using the Lynds-Worden-Harvey method have found no statistically significant surface detail.

McDonnell and Bates [1976] (see also Bates et al [1977 ∅∅IV,V]) contend that the Lynds-Worden-Harvey image (3.131) does not have the form of the true image but is in fact blurred by random errors in the estimates $\underset{\sim}{x}_{mi}$ (3.124) of the true positions $\underset{\sim}{x}'_{mi}$ (3.121) of the speckles in the speckle image. This means that (3.127) is not valid in practice. The delta functions in the first term on the last line of (3.126) are clustered in a region of nonzero extent, thereby broadening the first term of (3.128). The first term of (3.129) and the Lynds-Worden-Harvey images (3.130, 3.131) are similarly broadened. McDonnell and Bates model the blurring by a point spread function $h_L(\underset{\sim}{x})$, i.e.

$$s'_{LWH}(\underset{\sim}{x}) = A' f(\underset{\sim}{x}) \circledcirc h_D(\underset{\sim}{x}) \circledcirc h_L(\underset{\sim}{x}). \tag{3.132}$$

They note that $h_L(\underset{\sim}{x})$ is likely to depend on the form of $f(\underset{\sim}{x})$. They take Lynds et al's [1976] image of Betelgeuse and attempt to remove the blurring by assuming that $h_L(\underset{\sim}{x})$ is circularly symmetric, then separating $h_L(\underset{\sim}{x})$ from $f(\underset{\sim}{x})$ and $h_D(\underset{\sim}{x})$ by inspecting the zeros of the radial average of the visibility

$$S'_{LWH}(\underset{\sim}{u}) = A' F(\underset{\sim}{u}) H_D(\underset{\sim}{u}) H_L(\underset{\sim}{u}), \tag{3.133}$$

then Wiener filtering the estimate of $H_L(\underset{\sim}{u})$ from (3.133). The resulting enhanced image of Betelgeuse [McDonnell and Bates 1976, Bates et al 1977] has considerably more pronounced detail than the image of Lynds et al [1976].

Bakut et al [1984] describe a speckle imaging method which is a translation of the Lynds-Worden-Harvey method into the Fourier domain, with embellishments related to speckle interferometry (ϕ3.3.1) and the Knox-Thompson method (ϕ3.5.2). They claim that their method does not suffer from blurring due to the inaccurate estimation of speckle positions. Unfortunately this immunity arises not as a consequence of their method but rather from an implicit assumption in their analysis that the speckle position estimates are indeed perfect. When this assumption is removed their method has resolution equivalent to the Lynds-Worden-Harvey method.

von der Heide [1978] describes a speckle imaging method which can be considered an elaboration of Lynds et al's [1976] attempt to estimate $h_m(\underset{\sim}{x})$ from $s_m(\underset{\sim}{x})$. von der Heide solves for $h_m(\underset{\sim}{x})$ from $s_m(\underset{\sim}{x})$ with an iterative procedure incorporating a matrix formulation of least-squares minimisation. The method is not restricted to objects of very small angular extent and can incorporate a priori knowledge of the form of the object. von der Heide presents one-dimensional computer simulations of the method and notes that its major drawback is its requirement for an enormous amount of computation.

Bates and Milner [1978] describe an extension to the Lynds-Worden-Harvey method which allows objects (e.g. star clusters) consisting of a collection of unresolved parts to be imaged. They call their technique speckle masking (see also Bates et al [1978a], Bates et al [1978b], Milner and Bates [1979], and Milner [1979 ϕϕ6,7]), which is not to be confused with Weigelt et al's speckle masking (ϕ3.5.3). Bates and Milner's [1978] speckle masking incorporates elements of the Lynds-Worden-Harvey method, speckle interferometry, and the crystallographic heavy-atom methods mentioned in Section 3.4.1. Speckle masking begins by forming the image $s_{LWHm}(\underset{\sim}{x})$ described by (3.126). The function $d_m(\underset{\sim}{x})_e$ which is correlated with $s_m(\underset{\sim}{x})$ is appropriately called a speckle mask when the correlation is performed optically, hence the name "speckle masking".

Next, a threshold is lowered through $s_{LWHm}(\underset{\sim}{x})$ until the background is reached, and the positions and brightnesses of any parts of $s_{LWMm}(\underset{\sim}{x})$ exceeding the threshold are taken to form a basic star pattern. Bates and Milner argue that when one star in the star cluster $f(\underset{\sim}{x})$ is significantly brighter than all of the others, the brightest speckles in $s_m(\underset{\sim}{x})$ are more often than not associated with the brightest star in $f(\underset{\sim}{x})$. The central part of $s_{LWHm}(\underset{\sim}{x})$ then comprises mostly the true image of the brightest star, and true images of the other stars in $f(\underset{\sim}{x})$ appear with approximately their correct positions and brightnesses in $s_{LWHm}(\underset{\sim}{x})$. It is not necessary for all of the other stars to be recognisable above the background in $s_{LWHm}(\underset{\sim}{x})$, because Bates and Milner then compare the basic star pattern with the estimated autocorrelation $\mathbf{A}[f(\underset{\sim}{x})]_e$ from speckle interferometry (3.21) in a procedure which can unravel the form of $s_D(\underset{\sim}{x})$ from $\mathbf{A}[f(\underset{\sim}{x})]_e$ even if the basic star pattern is an incomplete or inaccurate representation of $f(\underset{\sim}{x})$. The comparison procedure is related to the algorithms described by Baldwin and Warner [1978a]. If the basic star pattern is significantly asymmetric, the mirror-image ambiguity in reconstructing $s_D(\underset{\sim}{x})$ from $\mathbf{A}[f(\underset{\sim}{x})]_e$ (cf. ∅3.4.2) is resolved.

3.6.2 SHIFT-AND-ADD

Bates and Cady [1980] describe a speckle imaging technique they call shift-and-add, in which the brightest point in each speckle image is located, the speckle image is translated or shifted so that the brightest point lies at the centre of image space, and the shifted speckle image is added to the other speckle images which have been similarly processed. This particularly simple technique is essentially a version of the Lynds-Worden-Harvey method without prefiltering (∅3.6.1) in which only the brightest speckle of each speckle image is chosen as a bright speckle. The motivation for shift-and-add came from considering situations in which there might actually be only one bright speckle in the speckle image, for instance when there are too few photons in the speckle image to adequately form any but the brightest speckles [Bates 1982b ∅8.7] or when the speckle image is formed with wideband light [Bates and Cady 1980] (cf. ∅2.5). A number of variations on Bates and Cady's technique have since been described which differ in details such as how the shifted speckle

images are weighted before they are added together and whether only the single brightest speckle or a few of the almost-brightest speckles [cf. Bagnuolo 1982] from each speckle image are shifted to the centre of image space. These variants all have similar characteristics and are here referred to collectively as basic shift-and-add methods. The original proposal of Bates and Cady [1980] is here called simple shift-and-add.

Using the notation with which the Lynds-Worden-Harvey method is described in Section 3.6.1, simple shift-and-add is characterised by a speckle mask

$$d_m(\underset{\sim}{x})_e = \delta(\underset{\sim}{x} - \underset{\sim}{x}_m) \tag{3.134}$$

containing one delta function of unit strength whose position $\underset{\sim}{x}_m$ is that of the brightest point in $s_m(\underset{\sim}{x})$. The simple shift-and-add image $s_{SA}(\underset{\sim}{x})$ is defined by

$$\begin{aligned} s_{SA}(\underset{\sim}{x}) &= \langle d_m(\underset{\sim}{x})_e \otimes s_m(\underset{\sim}{x}) \rangle_M \\ &= \langle s_m(\underset{\sim}{x} + \underset{\sim}{x}_m) \rangle_M \\ &= f(\underset{\sim}{x}) \odot \langle h_m(\underset{\sim}{x} + \underset{\sim}{x}_m) \rangle_M \\ &= f(\underset{\sim}{x}) \odot h_{SA}(\underset{\sim}{x}), \end{aligned} \tag{3.135}$$

where

$$h_{SA}(\underset{\sim}{x}) = \langle h_m(\underset{\sim}{x} + \underset{\sim}{x}_m) \rangle_M \tag{3.136}$$

is the simple shift-and-add point spread function.

The simple shift-and-add image, like the $s_{LWH}(\underset{\sim}{x})$ (3.129) formed by Lynds-Worden-Harvey processing without the prefiltering step, possesses a bright central peak and an extended background or fog [Bates and Cady 1980]. If $f(\underset{\sim}{x})$ happens to contain a single unresolved part (the reference part) that is considerably brighter than the rest of $f(\underset{\sim}{x})$, then $s_{LA}(\underset{\sim}{x})$ shows a diffraction-limited version of $f(\underset{\sim}{x})$ on top of the fog, with the bright central peak of $s_{SA}(\underset{\sim}{x})$ corresponding to the bright reference part of $f(\underset{\sim}{x})$ [Bates and Cady 1980, Cady and Bates 1980, Cady 1980 §§5,6, Bates 1982b §8.7, Hunt et

al 1983]. This version of $f(\underset{\sim}{x})$ remains diffraction-limited even when the imaging instrument is severely aberrated [Cady and Bates 1980, Bates et al 1981, Cady 1980 §§6.5,6.6]. When the reference part of $f(\underset{\sim}{x})$ is not considerably brighter than the rest of $f(\underset{\sim}{x})$, spurious details or ghosts [Bates and Cady 1980] appear in $s_{SA}(\underset{\sim}{x})$. The brightness of the ghosts compared to the brightness of the "true" detail in $s_{SA}(\underset{\sim}{x})$ increases as the reference part of $f(\underset{\sim}{x})$ becomes less bright with respect to the rest of $f(\underset{\sim}{x})$. Ghosting can be considered as a dependence of the form of $h_{SA}(\underset{\sim}{x})$ on the form of $f(\underset{\sim}{x})$. The simple shift-and-add method is illustrated in one dimension in Figure 3.9. As the number of independent speckle images averaged in (3.135) increases, the standard deviation of the average at each point in image space decreases in proportion to $1/\sqrt{M}$ and an image resembling the true image with ghosts and fog gradually appears (Figures 3.9c-h).

Another basic shift-and-add technique called adjusted shift-and-add is described by Bates et al [1981]. This differs from simple shift-and-add only in that each shifted speckle image is weighted by the brightness of its brightest point before being added to the others, so that

$$d_m(\underset{\sim}{x})_e = s_m(\underset{\sim}{x}_m)\,\delta(\underset{\sim}{x} - \underset{\sim}{x}_m) \tag{3.137}$$

and

$$\begin{aligned} s_{ASA}(\underset{\sim}{x}) &= \langle d_m(\underset{\sim}{x})_e \otimes s_m(\underset{\sim}{x})\rangle_M \\ &= \langle s_m(\underset{\sim}{x}_m)\; s_m(\underset{\sim}{x} + \underset{\sim}{x}_m)\rangle_M \\ &= f(\underset{\sim}{x}) \odot \langle s_m(\underset{\sim}{x}_m)\; h_m(\underset{\sim}{x} + \underset{\sim}{x}_m)\rangle_M . \end{aligned} \tag{3.138}$$

The brightness of the fog and the ghosts tends to be somewhat less in $s_{ASA}(\underset{\sim}{x})$ than in $s_{SA}(\underset{\sim}{x})$ [Bates et al 1981, Bates and Fright 1982] but the images are otherwise qualitatively similar.

The fog in basic shift-and-add images can be suppressed by forming the basic shift-and-add image of an unresolved reference object $f_o(\underset{\sim}{x}) \approx a\,\delta(\underset{\sim}{x})$ to give

$$s_{oSA}(\underset{\sim}{x}) = f_o(\underset{\sim}{x}) \odot h_{oSA}(\underset{\sim}{x}) \approx a\,h_{oSA}(\underset{\sim}{x}), \tag{3.139}$$

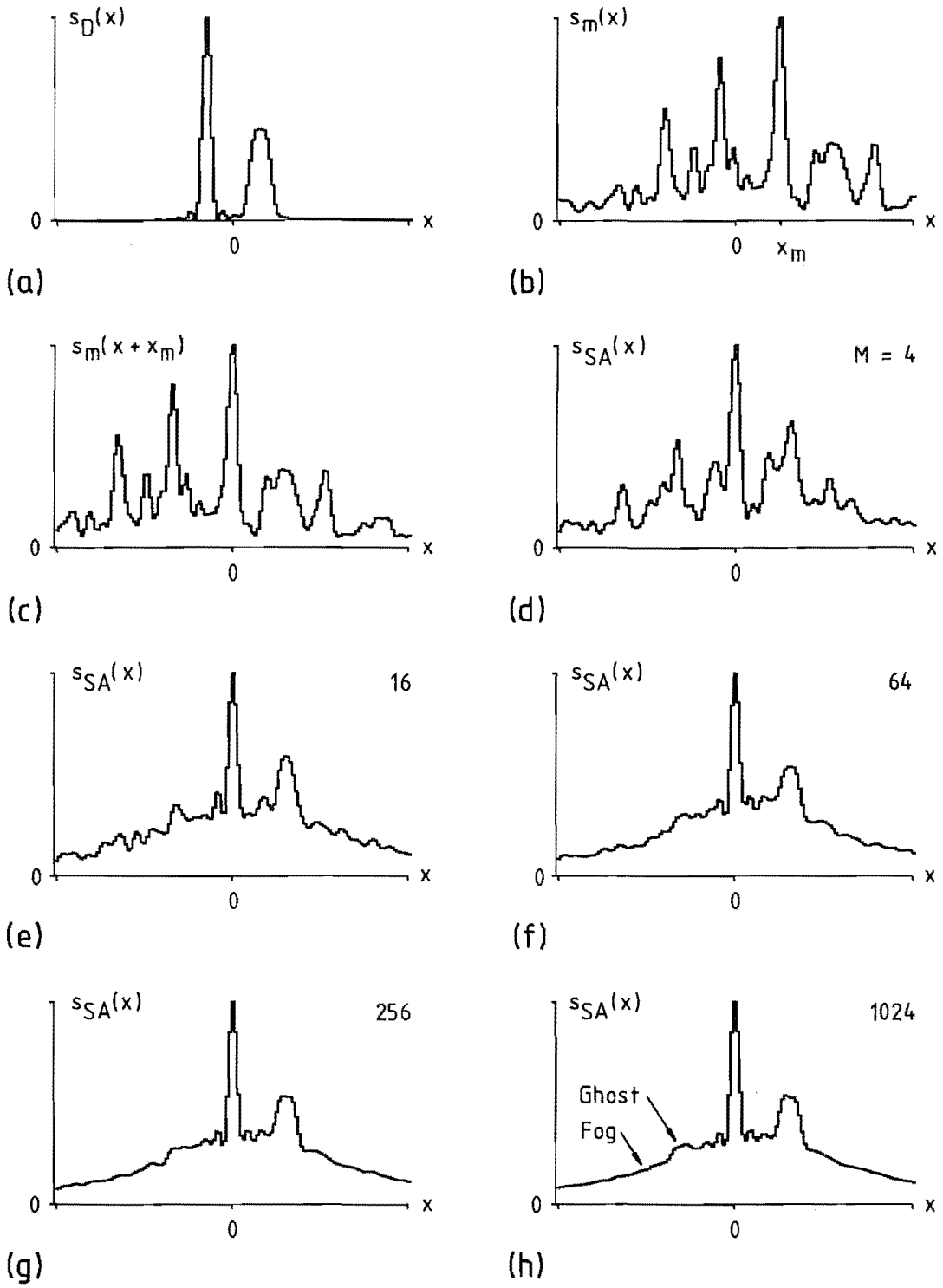


Figure 3.9: One-dimensional computer simulations illustrating simple shift-and-add.

(a) True image.

(b) Typical speckle image, whose brightest pixel occurs at $x = x_m$.

(c) - (h) Simple shift-and-add images

$s_{SA}(x) = \langle s_m(x + x_m) \rangle_M$ for M = 1 (c), 4 (d), 16 (e), 64 (f), 256 (g), 1024 (h) speckle images.

then deconvolving $s_{oSA}(\underset{\sim}{x})$ from $s_{SA}(\underset{\sim}{x})$ by a technique such as Wiener filtering (cf. 3.19) or CLEAN [Högbom 1974] [cf. Bates et al 1984, Bates and McDonnell 198_ ϕ16], then convolving the result with the diffraction-limited psf of the telescope. The resulting defogged image is given by

$$\begin{aligned} s'_{SA}(\underset{\sim}{x}) &= s_{SA}(\underset{\sim}{x}) \odot^{-1} s_{oSA}(\underset{\sim}{x}) \odot h_D(\underset{\sim}{x}) \\ &= f(\underset{\sim}{x}) \odot h_D(\underset{\sim}{x}) \odot h_{SA}(\underset{\sim}{x}) \odot^{-1} a\, h_{oSA}(\underset{\sim}{x}) \\ &= \frac{1}{a} s_D(\underset{\sim}{x}) \odot \left[h_{SA}(\underset{\sim}{x}) \odot^{-1} h_{oSA}(\underset{\sim}{x})\right]. \end{aligned} \tag{3.140}$$

The object-dependence of $h_{SA}(\underset{\sim}{x})$ means that deconvolution by $h_{oSA}(\underset{\sim}{x})$ does not remove ghosts, but the fog is largely removed (see Figures 3.10a,b). An alternative approach to defogging is to estimate the fog and subtract the estimate from the shift-and-add image. Bagnuolo [1985a] estimates the fog by forming an image comprising the sum of delta functions at the location of each speckle image used to determine the shifts in shift-and-add, i.e.

$$d(\underset{\sim}{x}) = \langle d_m(\underset{\sim}{x})_e \rangle_M \tag{3.141}$$

(cf. 3.134, 3.137), and then cross-correlating this image with the unshifted ensemble average $\langle s_m(\underset{\sim}{x}) \rangle_M$ (which equals the long exposure image (3.10) in the limit of $M \to \infty$) to form the estimated fog

$$s_f(\underset{\sim}{x}) = d(\underset{\sim}{x}) \otimes \langle s_m(\underset{\sim}{x}) \rangle_M . \tag{3.142}$$

The defogged shift-and-add image is given by

$$s'_{SA}(\underset{\sim}{x}) = s_{SA}(\underset{\sim}{x}) - s_f(\underset{\sim}{x}) \tag{3.143}$$

(see Figures 3.10c,d). This procedure is reminiscent of Welter and Worden's [1978] method or long exposure subtraction in speckle interferometry (ϕ3.3.2) and evidently (Figure 3.10d) causes a similar attenuation of spatial frequencies below the seeing limit.

Hunt et al [1983] have confirmed theoretically, effectively under the assumption of $D >> r_o$, that the simple shift-and-add psf contains a diffraction-limited component if the object possess spatial

frequencies out to the diffraction limit [cf. Freeman et al 1985]. Analytical descriptions of the effects of ghosting are presented by Bagnuolo [1982] and Hunt et al [1983 §4] for objects with unresolved parts (e.g. binary stars) and by Hunt et al [1983 §5] and Bagnuolo [1985a, 1985b] for extended objects. Bagnuolo [1984] analyses the effects of nonisoplanatism on basic shift-and-add (cf. §5.3) using a single phase screen model of the atmosphere (§2.7).

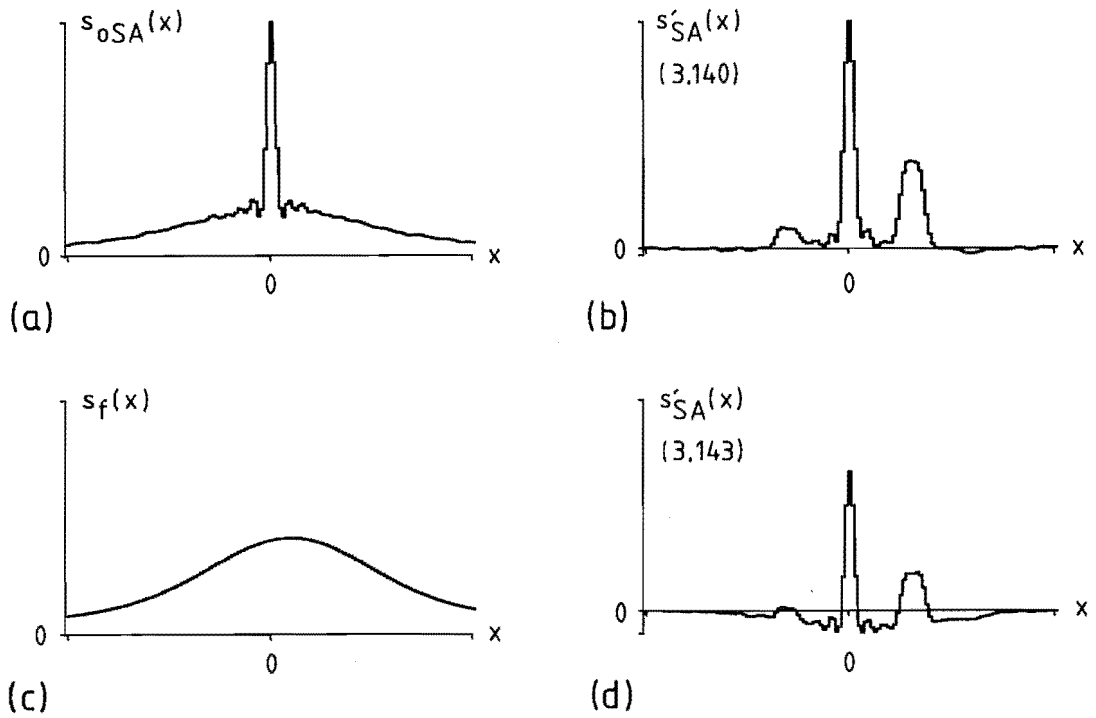


Figure 3.10: One-dimensional computer simulations illustrating two ways of removing the fog from the simple shift-and-add image: by deconvolution (a, b), or by subtraction (c, d). 1000 speckle images were used to form (a) and (c).

(a) Simple shift-and-add image of a point reference object.

(b) The result of deconvolving (a) from Figure 3.9h by Wiener filtering (cf. 3.19) and then convolving by the diffraction-limited transfer function.

(c) Bagnuolo's [1985a] estimate of the fog (3.142), displayed on the same vertical scale as Figure 3.9h.

(d) The result of subtracting (c) from Figure 3.9h, displayed on the same vertical scale as Figure 3.9h.

Astronomical applications of basic shift-and-add have been reported by Bagnuolo and McAlister [1983], Hutter et al [1985], Baba et al [1985], and Bagnuolo [1985a, 1985b]. The first three of these papers describe observations of binary stars. Shift-and-add is able to resolve the 180° ambiguity of position angle inherent in speckle interferometric observations (cf. ϕ3.4.2). The two papers by Bagnuolo [1985a, 1985b] present an image of Betelgeuse which suggests the presence of an irregular envelope of material surrounding the star [cf. Roddier and Roddier 1985].

Basic shift-and-add has been applied to ultrasonic imaging through an inhomogeneous medium by Bates and Robinson [1981] (see also Bates et al [1982] and Bates and Minard [1983, 1984]). In this case the object $f_c(\underset{\sim}{x})$ is spatially coherent and the speckle images are complex amplitudes (cf. ϕ1.6.3). The simple shift-and-add image defined by

$$s_{SA}(\underset{\sim}{x}) = \langle s_m(\underset{\sim}{x} + \underset{\sim}{x}_m)\, e^{-j\,\mathrm{phase}[s_m(\underset{\sim}{x}_m)]} \rangle_M$$

$$= f_c(\underset{\sim}{x}) \odot \langle h_m(\underset{\sim}{x} + \underset{\sim}{x}_m)\, e^{-j\,\mathrm{phase}[s_m(\underset{\sim}{x}_m)]} \rangle_M \qquad (3.144)$$

is qualitatively similar to the simple shift-add-image in the spatially incoherent case (3.135), including the formation of ghosts, except that the random background averages towards zero instead of towards a positive-valued fog. Statistically independent speckle images are obtained by forming the images with quasi-monochromatic ultrasound of different mean frequencies [Bates and Robinson 1981].

An extension to basic shift-and-add in ultrasonic imaging is described by Bates and Robinson [1982]. In this technique, called stochastic shift-and-add, synthetically distorted quasi-speckle images $s_{mn}(\underset{\sim}{x})$ are generated from each $s_m(\underset{\sim}{x})$ by applying a computer-generated pseudo-random phase distortion $\phi_{mn}(\underset{\sim}{u})$ to $S_m(\underset{\sim}{u})$, giving

$$s_{mn}(\underset{\sim}{x}) \leftrightarrow S_{mn}(\underset{\sim}{u}) = S_m(\underset{\sim}{u})\, e^{-j\phi_{mn}(\underset{\sim}{u})}. \qquad (3.145)$$

Simple shift-and-add is then performed over all $s_{mn}(\underset{\sim}{x})$ from all $s_m(\underset{\sim}{x})$, forming the stochastic shift-and-add image

$$s_{SSA}(\underset{\sim}{x}) = \langle s_{mn}(\underset{\sim}{x} + \underset{\sim}{x}_{mn})\, e^{-j\,\mathrm{phase}[s_{mn}(\underset{\sim}{x}_{mn})]} \rangle_{MN}$$

$$= f(\underset{\sim}{x}) \odot \langle h_{mn}(\underset{\sim}{x} + \underset{\sim}{x}_{mn})\, e^{-j\,\mathrm{phase}[s_{mn}(\underset{\sim}{x}_{mn})]} \rangle_{MN}. \qquad (3.146)$$

The stochastic shift-and-add image is often a more faithful rendition of the true image than is the simple shift-and-add image (3.144) when the object consists of a collection of unresolved parts, e.g. microcalcifications in human breast tissue [cf. Robinson and Bates 1980]. It can even be a good likeness to the true image when only a single speckle image is used in (3.146) [Bates and Robinson 1982, Bates 1982b ϕ12.3]. Bates [1976] demonstrates the latter possibility in the context of spatially incoherent speckle imaging.

The main advantage of simple shift-and-add over the other speckle imaging techniques described in this chapter is that it is operationally very simple. Just as speckle interferometry (ϕ3.3.1) is well suited to coherent optical computation, so is simple shift-and-add well suited to digital computation because it involves only elementary image domain operations (compare, shift, add) that can be rapidly performed by a microcomputer [cf. Bates et al 1981]. The other basic shift-and-add techniques described above require more computational effort. For objects whose basic shift-and-add images are so badly contaminated by ghosts that they are not useful as is, basic shift-and-add can usefully be incorporated into composite speckle imaging schemes [Bates and Fright 1982] with other techniques such as speckle interferometry (ϕ3.3) and Fienup's algorithms (ϕ3.4.2). An important property of all the basic shift-and-add images (3.135, 3.138, 3.140, 3.143, 3.144, 3.146) is that they are each convolutions of $f(\underset{\sim}{x})$ with a point spread function if the speckle images themselves are isoplanatic (3.3e). This property may be exploitable by recently developed computational procedures employing complex visibility zeros (ϕ3.5.4) [cf. Lane et al 198_].

3.6.3 RECENT VARIANTS OF SHIFT-AND-ADD

In recent years, several shift-and-add techniques significantly more intricate than basic shift-and-add have been

developed. This subsection describes these new techniques.

Minard et al [1985] present an iterative procedure called correlation shift-and-add which can form faithful images of certain types of object, such as objects consisting of a collection of unresolved parts of comparable brightnesses (e.g. star clusters), whose basic shift-and-add images are severely contaminated by ghosts (∅3.6.2). Correlation shift-and-add is one of a number of extensions to basic shift-and-add studied by Minard [1985 ∅∅7-9] [cf. Bates et al 1985, Sinton et al 198_] in which a transformation is applied to each speckle image in the ensemble $\{s_m(\underset{\sim}{x})\}$ in order to make the position of the brightest point in each $h_m(\underset{\sim}{x})$ more accurately identifiable. In correlation shift-and-add, each speckle image is correlated by an estimate $f(\underset{\sim}{x})_e$ of $f(\underset{\sim}{x})$ to form an ensemble of modified speckle images

$$s'_m(\underset{\sim}{x}) = f(\underset{\sim}{x})_e \otimes s_m(\underset{\sim}{x}) = \left[f(\underset{\sim}{x})_e \otimes f(\underset{\sim}{x})\right] \odot h_m(\underset{\sim}{x}). \tag{3.147}$$

When $f(\underset{\sim}{x})_e$ is a good estimate of $f(\underset{\sim}{x})$, i.e.

$$f(\underset{\sim}{x})_e \approx f(\underset{\sim}{x}), \tag{3.148}$$

(3.147) becomes

$$s'_m(\underset{\sim}{x}) \approx \mathbf{A}\left[f(\underset{\sim}{x})\right] \odot h_m(\underset{\sim}{x}). \tag{3.149}$$

Ghosting in basic shift-and-add (∅3.6.2) occurs when the brightest point in $s_m(\underset{\sim}{x})$ does not correspond to the brightest point in $h_m(\underset{\sim}{x})$ in a significant proportion of the speckle images. If (as is often the case) $f(\underset{\sim}{x})_e \otimes f(\underset{\sim}{x})$ is more highly peaked than $f(\underset{\sim}{x})$, for instance when $f(\underset{\sim}{x})$ consists of an irregularly spaced collection of unresolved parts, then the position $\underset{\sim}{x}'_m$ of the brightest point in each $s'_m(\underset{\sim}{x})$ is more likely to be equal to the position of the brightest point in $h_m(\underset{\sim}{x})$ than is the position $\underset{\sim}{x}_m$ of the brightest point in $s_m(\underset{\sim}{x})$. Any of the basic shift-and-add algorithms of Section 3.6.2 can then be advantageously adapted so that each unmodified speckle image is shifted according to $\underset{\sim}{x}'_m$ instead of $\underset{\sim}{x}_m$ before being added. The resulting modified basic shift-and-add image exhibits less ghosting than the basic shift-and-add image formed with shifts according to $\underset{\sim}{x}_m$.

In the discussion which follows, simple shift-and-add (3.135) is used as a representative example of basic shift-and-add. The modified basic shift-and-add image is then given by

$$s'_{SA}(\underset{\sim}{x}) = \langle s_m(\underset{\sim}{x} + \underset{\sim}{x}'_m)\rangle_M = f(\underset{\sim}{x}) \circledcirc \langle h_m(\underset{\sim}{x} + \underset{\sim}{x}'_m)\rangle_M \,. \qquad (3.150)$$

The form of the object is unknown a priori, so correlation shift-and-add forms and iteratively refines an estimate $f(\underset{\sim}{x})_e$ of $f(\underset{\sim}{x})$ with the following algorithm:

(a) Estimate the extent of $f(\underset{\sim}{x})$ by performing speckle interferometry to obtain $\mathbf{A}[f(\underset{\sim}{x})]_e$ (3.21, 3.22) and setting an extent estimate $\underset{\sim}{E}_e$ equal to half of the effective extent of $\mathbf{A}[f(\underset{\sim}{x})]_e$. Construct a rectangular window function $w(\underset{\sim}{x})$ with a width perhaps 10% greater than $\underset{\sim}{E}_e$.

(b) To begin the iteration, set $\{s'_m(\underset{\sim}{x})\}$ equal to $\{s_m(\underset{\sim}{x})\}$.

(c) Form the modified basic shift-and-add image $s'_{SA}(\underset{\sim}{x})$ according to (3.150).

(d) Slide the window function across $s'_{SA}(\underset{\sim}{x})$ and find the position $\underset{\sim}{x} = \underset{\sim}{x}_w$ for which $\int |s'_{SA}(\underset{\sim}{x})\, w(\underset{\sim}{x} - \underset{\sim}{x}_w)| \,.\, d\underset{\sim}{x}$ is maximised. Set

$$f(\underset{\sim}{x})_e = s_{SA}(\underset{\sim}{x})\, w(\underset{\sim}{x} - \underset{\sim}{x}_w). \qquad (3.151)$$

(e) Form a new ensemble $\{s'_m(\underset{\sim}{x})\}$ of modified speckle images according to (3.147).

(f) Repeat steps (c) - (e) until successive versions of $s'_{SA}(\underset{\sim}{x})$ differ by less than a prescribed threshold. The correlation shift-and-add image $s_{CAA}(\underset{\sim}{x})$ is then set equal to the most recently obtained $s'_{SA}(\underset{\sim}{x})$ [Minard et al 1985] or alternatively to the most recently obtained $s'_{SA}(\underset{\sim}{x})\, w(\underset{\sim}{x} - \underset{\sim}{x}_w)$ [Sinton et al 198_].

Minard et al [1985] and Minard [1985] apply correlation shift-and-add to ultrasonic imaging, for which the basic shift-and-add image contains no fog (§3.6.2). To adapt correlation shift-and-add to astronomical speckle imaging the fog must be removed from $s'_{SA}(\underset{\sim}{x})$ or

from $f(\underset{\sim}{x})_e$ in step (d) above [cf. Sinton et al 198_].

The correlation (3.147) is equivalent to the convolution

$$s'_m(\underset{\sim}{x}) = g(\underset{\sim}{x})_e \odot s_m(\underset{\sim}{x}) \tag{3.152}$$

where

$$g(\underset{\sim}{x})_e = f^*(-\underset{\sim}{x})_e \tag{3.153}$$

is an estimate of the matched filter for $f(\underset{\sim}{x})$. This is the essence of the matched-filter shift-and-add proposed independently of Minard et al [1985] by Ribak et al [1984, 1985a] and Ribak [198_]. Matched-filter shift-and-add is closely related to correlation shift-and-add. It also aims to accurately determine the locations of local maxima of $h_m(\underset{\sim}{x})$ from $s_m(\underset{\sim}{x})$. It also begins with an initial estimate of $f(\underset{\sim}{x})$ and iteratively refines the estimate from $\{s_m(\underset{\sim}{x})\}$ [Ribak 198_]. The identity between (3.152) and (3.147) describes the essential identity between the two methods.

Bagnuolo [1985a, 1985b] describes an iterative algorithm which promises to be able to suppress ghosting in the basic shift-and-add image of an arbitrary object. From a model of the first-order statistics of speckle intensity (∅2.5.1), Bagnuolo derives an expression for $h_{SA}(\underset{\sim}{x})$ in terms of $f(\underset{\sim}{x})$ [cf. Hunt et al 1983 ∅6]. This expression allows the shift-and-add image of a given object to be computed theoretically. The iterative algorithm performs the inverse operation of determining $f(\underset{\sim}{x})$ from $s_{SA}(\underset{\sim}{x})$. It is similar in some ways to Fienup's algorithms (∅3.4.2), and has the following steps:

(a) Form an initial estimate $f(\underset{\sim}{x})_e$ of $f(\underset{\sim}{x})$.

(b) Compute $h_{SA}(\underset{\sim}{x})_e$ from $f(\underset{\sim}{x})_e$ using Bagnuolo's [1985a, 1985b] theoretical results.

(c) Deconvolve $h_{SA}(\underset{\sim}{x})_e$ from the known shift-and-add image $s_{SA}(\underset{\sim}{x})$ to give

$$f'(\underset{\sim}{x})_e = s_{SA}(\underset{\sim}{x}) \odot^{-1} h_{SA}(\underset{\sim}{x})_e \ . \tag{3.154}$$

(d) Form a new $f(\underset{\sim}{x})_e$ by imposing constraints upon $f'(\underset{\sim}{x})_e$:

$$f(\underset{\sim}{x})_e = \begin{cases} f'(\underset{\sim}{x})_e & \text{if } \underset{\sim}{x} : f'(\underset{\sim}{x})_e \geq 0 \text{ and inside estimated extent of } f(\underset{\sim}{x}) \\ 0 & \text{otherwise} \end{cases} \quad (3.155)$$

(e) Repeat steps (b) - (d) until convergence is achieved, leaving $f(\underset{\sim}{x})_e$ as the estimate of $f(\underset{\sim}{x})$.

Bagnuolo [1985a, 1985b] reports that the algorithm converges rapidly for objects containing bright regions of small extent (e.g. a barely resolved star). He presents images from computer simulations and an image of Betelgeuse formed from astronomical data. The image of Betelgeuse shows a bright central disc surrounded by an extended region of irregular detail. The image is qualitatively similar to the images formed by Roddier and Roddier [1985] and Bates et al [198_] (cf. ϕ3.4.2).

Christou et al [1984, 1985, 1986, 198_] (see also Christou [1985a ϕ6]) describe a close relative of the Lynds-Worden-Harvey method (ϕ3.6.1) which, following Christou et al [1984], is here called weighted shift-and-add. The objective of weighted shift-and-add is to produce an image that is representative of the true image $s_D(\underset{\sim}{x})$ and whose form does not depend on the parameters of the seeing [Christou 1985a], i.e. the image contains no background fog (cf. ϕϕ 3.6.1, 3.6.2). Instead of prefiltering individual speckle images as the Lynds-Worden-Harvey method does, weighted shift-and-add forms a basic shift-and-add image (cf. ϕ3.6.2) and filters the fog from that. Weighted shift-and-add begins by forming a speckle mask (ϕ3.6.1) $d_m(\underset{\sim}{x})_e$ containing delta functions whose positions and strengths are equal to the positions and brightnesses of the local maxima of $s_m(\underset{\sim}{x})$, i.e.

$$d_m(\underset{\sim}{x})_e = \sum_{i=1}^{I_m} s_m(\underset{\sim}{x}_{mi})\, \delta(\underset{\sim}{x} - \underset{\sim}{x}_{mi}). \quad (3.156)$$

Every local maximum of $s_m(\underset{\sim}{x})$ above a threshold level representing the estimated noise level in $s_m(\underset{\sim}{x})$ is included into $d_m(\underset{\sim}{x})_e$. The threshold level is typically 5% of the maximum brightness of $s_m(\underset{\sim}{x})$ [Christou

1985a ϕ6.4]. The weighted shift-and-add image $s_{WSA}(\underset{\sim}{x}) \leftrightarrow S_{WSA}(\underset{\sim}{u})$ is then computed according to

$$S_{WSA}(\underset{\sim}{u}) = \frac{< \frac{S_m(\underset{\sim}{u})}{D_m(\underset{\sim}{u})_e} \cdot |D_m(\underset{\sim}{u})_e|^2 >_M}{<|D_m(\underset{\sim}{u})_e|^2>_M}$$

$$= \frac{<S_m(\underset{\sim}{u})\ D^*_m(\underset{\sim}{u})_e>_M}{<D_m(\underset{\sim}{u})_e\ D^*_m(\underset{\sim}{u})_e>_M} . \qquad (3.157)$$

In the image domain (3.157) becomes

$$s_{WSA}(\underset{\sim}{x}) = <d_m(\underset{\sim}{x})_e \otimes s_m(\underset{\sim}{x})>_M \odot^{-1} <d_m(\underset{\sim}{x})_e \otimes d_m(\underset{\sim}{x})_e>_M , \qquad (3.158)$$

where $\odot^{-1}$ represents the deconvolution accomplished by dividing by $<D_m(\underset{\sim}{u})_e\ D^*_m(\underset{\sim}{u})_e>_M$ in the Fourier domain (cf. 3.18). Christou et al [1985] note that the division is computationally stable because $<D_m(\underset{\sim}{u})_e\ D^*_m(\underset{\sim}{u})_e>_M = <|D_m(\underset{\sim}{u})|^2_e>_M$ is greater than zero for all $\underset{\sim}{u}$. The first term on the right hand side of (3.158) is exactly equivalent to the Lynds-Worden-Harvey image without prefiltering (3.129). The second term of the right hand side of (3.158) is the averaged autocorrelation of the speckle masks, which provides a synthetic estimate of the fog. Deconvolving the second quantity from the first suppresses the fog almost completely, as the stellar images presented by Christou et al [1985, 1986, 198_] and Christou [1985a] demonstrate by example.

The success of weighted shift-and-add in forming a faithful, fog-free estimate of the true image is dependent upon the same conditions as is the success of the Lynds-Worden-Harvey method (ϕ3.6.1), and therefore will probably be applicable only to barely resolved bright objects unless the formation of the speckle mask is augmented by techniques such as matched filtering [Ribak et al 1984, 1985a, Ribak 198_, Christou et al 198_]. The conditions required for weighted shift-and-add to successfully remove the fog can be established by inserting Lynds et al's [1976] model of the speckle point spread function (3.122) into (3.158) (cf. 3.3e). This gives

$$s_{WSA}(\underset{\sim}{x}) = \langle d_m(\underset{\sim}{x})_e \otimes [f(\underset{\sim}{x}) \odot (h_D(\underset{\sim}{x}) \odot d_m(\underset{\sim}{x}) + n_m(\underset{\sim}{x}))] \rangle_M$$
$$\odot^{-1} \langle d_m(\underset{\sim}{x})_e \otimes d_m(\underset{\sim}{x})_e \rangle_M$$
$$= f(\underset{\sim}{x}) \odot h_D(\underset{\sim}{x}) \odot [\langle d_m(\underset{\sim}{x})_e \otimes d_m(\underset{\sim}{x}) \rangle_M + \langle d_m(\underset{\sim}{x})_e \otimes n_m(\underset{\sim}{x}) \rangle_M]$$
$$\odot^{-1} \langle d_m(\underset{\sim}{x})_e \otimes d_m(\underset{\sim}{x})_e \rangle_M$$
$$= s_D(\underset{\sim}{x}) \odot [\langle d_m(\underset{\sim}{x})_e \otimes d_m(\underset{\sim}{x}) \rangle_M + \langle d_m(\underset{\sim}{x})_e \otimes n_m(\underset{\sim}{x}) \rangle_M]$$
$$\odot^{-1} \langle d_m(\underset{\sim}{x})_e \otimes d_m(\underset{\sim}{x})_e \rangle_M . \quad (3.159)$$

Only if

$$d_m(\underset{\sim}{x})_e \approx d_m(\underset{\sim}{x}) \quad (3.160a)$$

and

$$\langle d_m(\underset{\sim}{x})_e \otimes n_m(\underset{\sim}{x}) \rangle_M \approx 0 \quad (3.160b)$$

will the fog be well removed, since then

$$s_{WSA}(\underset{\sim}{x}) \approx s_D(\underset{\sim}{x}) \odot [\langle d_m(\underset{\sim}{x})_e \otimes d_m(\underset{\sim}{x})_e \rangle_M \odot^{-1} \langle d_m(\underset{\sim}{x})_e \otimes d_m(\underset{\sim}{x})_e \rangle_M]$$
$$= s_D(\underset{\sim}{x}) \odot \delta(\underset{\sim}{x})$$
$$= s_D(\underset{\sim}{x}). \quad (3.161)$$

Condition (3.160a) requires that all speckles in $h_m(\underset{\sim}{x})$ be included into $d_m(\underset{\sim}{x})_e$, i.e. the threshold level mentioned above should be as low as possible and the detection of speckles in $s_m(\underset{\sim}{x})$ must be reliable. Condition (3.160b) expresses the intuitive assumption (stated as Equation 1 in Christou et al [1985]) that the average speckle of an unresolved object has the form of the diffraction-limited point spread function of the telescope, as has been verified theoretically by Hunt et al [1983] and Freeman et al [1985].

3.7 SUMMARY

Speckle images in optical astronomy are short exposure,

usually narrow band images of celestial objects, normally observed with conventional telescopes of large aperture (e.g. pupil diameter > 1m). Speckle processing techniques are data processing techniques that operate on ensembles of speckle images suffering statistically independent distortions due to atmospheric seeing. The intent of speckle processing is to obtain spatial resolution beyond the seeing limit imposed by the atmosphere on long exposure images. The original and most widely applied speckle processing technique is Labeyrie's speckle interferometry, which allows a diffraction-limited estimate of the autocorrelation of a celestial object to be formed despite atmospheric seeing, and also despite aberrations in the telescope optics provided the random distortions due to the atmosphere are worse than the fixed distortions due to the telescope. Estimating the autocorrelation of an object corresponds to estimating the squared magnitude of its visibility, in which the Fourier phase is lost. Certain characteristics of the object such as its size and approximate shape can be deduced from its autocorrelation, but the form of the true image of the object cannot be immediately recovered from its autocorrelation except for the special forms of object which can be imaged by speckle holography and related techniques. Fourier phase retrieval algorithms can be applied for objects of arbitrary form to reconstruct the object visibility phase discarded by speckle interferometry, which corresponds to determining the form of the true image from its diffraction-limited autocorrelation. Fourier phase retrieval tends to be computationally expensive.

Various speckle imaging techniques have been devised which attempt to preserve the object visibility phase information present in distorted form in speckle images. The different techniques have different strengths and weaknesses, producing estimates of the true image of varying quality for varying forms of object, quantities of detected light, and seeing conditions. An estimated true image which is insufficiently faithful to be useful on its own can be used to ease the computational burden and resolve fundamental ambiguities of Fourier phase retrieval. This chapter divides speckle imaging techniques broadly into Fourier domain methods and image domain methods. The image domain methods, typified by shift-and-add operations in image space, can be very simple conceptually and computationally. However their use of specific information from

individual speckle images, i.e. the positions of bright speckles, makes them difficult to describe analytically so that there are far fewer theoretical results for the image domain methods than for the Fourier domain methods.

Speckle interferometry and most speckle imaging techniques require the compensation of a transfer function or point spread function whose form depends on the statistical properties of the atmospheric seeing. The difficulty of this compensation hinders the quantitative measurement of intensity or brightness distributions of celestial objects by speckle processing. Positional information such as size, separation, or shape is usually more easily obtained. The standard calibration technique is to observe an unresolved reference object through seeing statistically similar to that through which the object of interest is observed. Much effort has been expended on trying to avoid the need for a reference object and on developing speckle processing techniques which do not require calibration for the seeing.

Astronomical applications of the speckle processing techniques described in this chapter have been reported for speckle interferometry (§3.3), speckle holography (§3.4.1), the Knox-Thompson method (§3.5.2), speckle masking (triple correlation) (§3.5.3), and members of each of the three divisions of shift-and-add techniques (§§ 3.6.1, 3.6.2, 3.6.3).

CHAPTER 4

COMPUTER SIMULATION OF ONE-DIMENSIONAL SPECKLE IMAGES

4.1 INTRODUCTION

Digital computer simulations of astronomical speckle processing are invoked extensively in this thesis. They provide illustrations for Chapters 2 and 3 and are the foundation of the computational investigations reported in Chapters 5 and 6. They are much used partly because computer simulated speckle images are easier to obtain than measured speckle images of astronomical objects, and partly because the ability to easily and repeatably control the statistical properties of the pseudo-random simulated atmosphere is a major asset in experimentation. The algorithms used to generate the computer simulated speckle images are described in this chapter. Section 4.2 explains why one-dimensional rather than two-dimensional speckle images have been simulated. Section 4.3 describes the computer representation of the image domain and visibility domain quantities present in the optical astronomical imaging system (Chapter 2) or manipulated by astronomical speckle processing techniques (Chapter 3). Section 4.4 describes the algorithm for simulating speckle images with perfect isoplanatism and no photon limiting. The modification of this algorithm to embrace nonisoplanatism is described in Section 4.5. Simulation of photon limiting is described in Section 4.6. Section 4.7 presents a concise summary of the parameters controlling the characteristics of the simulated speckle images.

4.2 WHY ONE DIMENSION

Speckle processing in optical astronomy is normally performed with two-dimensional speckle images. One-dimensional processing is common only in the infrared for which two-dimensional detectors are not readily available [Dainty 1984 φ7.5.2; cf. Sibille et al 1979, Dyck and Howell 1985]. However, computer simulation of two-dimensional operations is much more computationally expensive than

computer simulation of one-dimensional operations, for a given number of pixels per dimension, when the number of pixels per dimension is large enough to be useful. The computer simulations of speckle processing reported in this thesis have therefore been carried out in one dimension to reduce the computational burden involved.

The Fourier optical description of the optical astronomical imaging system and the formation of speckle images presented in Chapter 2 is applicable in both one and two dimensions. Furthermore, shift-and-add is equally applicable in one and two dimensions because its component operations of searching for brightest points, translating in image space, and adding images together can be applied essentially without change in any number of dimensions. The same is true of speckle interferometry. Previous experience with optical laboratory and digital computer simulations of shift-and-add has shown that its behaviour is qualitatively similar in one and two dimensions [cf. Bates and Fright 1982, Bates and Cady 1980]. All of this provides confidence that the conclusions drawn from the one-dimensional computational studies of shift-and-add and speckle interferometry in Chapter 5 also apply to two dimensions.

The zero-and-add speckle imaging technique introduced in Chapter 6 directly operates only on one-dimensional speckle images. Simulating speckle images in one dimension is therefore entirely appropriate to computational studies of zero-and-add. (Note that two-dimensional objects can be imaged by zero-and-add by performing zero-and-add on one-dimensional projections of two-dimensional speckle images, cf. Section 6.6.)

4.3 REPRESENTING THE IMAGING SYSTEM IN THE COMPUTER

The optical astronomical imaging system is simulated in the computer with data structures [cf. Wirth 1976] representing the object, the image, and related image domain and visibility domain quantities, and by algorithms implementing the relationships introduced in Chapters 1 and 2 (cf. 3.3) between these quantities. The data structures representing the image domain quantities (e.g. $f(x)$, $h_m(x)$, $s_m(x)$) and visibility domain quantities (e.g. $F(u)$,

$H_m(u)$, $S_m(u)$) are arrays of complex numbers which are the values of regularly spaced samples or pixels of the quantities in the real-world. Fourier transform operations between the image and visibility domains are performed by the Fast Fourier Transform (FFT) implementation [Cooley and Tukey 1965, Bergland 1969, Brigham 1974] of the discrete Fourier transform (DFT) [Bracewell 1978 φ18].

The image domain and visibility domain position variables x and u are embodied in the computer as indices n_x and n_u to the image and visibility domain arrays respectively. n_x and n_u are integers ranging from 1 to N_x and 1 to N_u respectively where N_x and N_u are the sizes of the image domain and visibility domain arrays. The particular implementation of the FFT used in the computer program [Ahmed and Rao 1975 p.79] constrains N_x and N_u to be equal to each other and to an integer power of 2, i.e.

$$N_x = N_u = SIZE = 2^i , \quad i \in I \tag{4.1}$$

(cf. φ4.7). The integer indices n_x, n_u in the computer are related to the real-world positions x, u by

$$x_n = \left(n_x - 1 - \frac{SIZE}{2}\right) \Delta x, \tag{4.2a}$$

$$u_n = \left(n_u - 1 - \frac{SIZE}{2}\right) \frac{SIZE}{\Delta x} \tag{4.2b}$$

where x has units of length in the image plane of the telescope, u has units of cycles per unit length, and Δx is the pixel or sample spacing in the image domain. So far as programming the simulations is concerned, Δx is arbitrary and does not appear explicitly in the computer program. The spatial scaling factors $-\ell$ between object space position α and image space position x (2.1), and $\bar{\lambda}\ell$ between pupil plane position ξ and spatial frequency u (3.3a), are likewise arbitrary when programming the simulations. In the computer they are normalised to the values

$$-\ell = 1, \tag{4.3a}$$

$$\bar{\lambda}\ell = 1, \tag{4.3b}$$

so that

$$\alpha = x \tag{4.4a}$$

(cf. 2.1) and

$$\xi = u \tag{4.4b}$$

(cf. 3.3a). These normalisations are not invoked in the equations in the following subsections of this chapter describing the computer algorithms, for the sake of preserving compatibility with the notation used elsewhere in this thesis.

In the remainder of this chapter and elsewhere in this thesis, computer simulated images and visibilities are written as functions of continuous position x or u instead of discrete position x_n or u_n (4.2a,b) except where the discreteness of the simulated position variables is a crucial part of the discussion (e.g. 4.8, 4.9, 4.13). This has been done to maintain notational compatibility with the discussions of Chapters 1 to 3. It should nevertheless be remembered that the simulated images are functions of discrete position and that the discrete Fourier transform rather than the continuous Fourier transform relates the image and visibility domains in the computer.

4.4 SIMULATING BRIGHT, ISOPLANATIC SPECKLE IMAGES

Speckle images with perfect isoplanatism and without photon limiting are simulated by convolving an object with a pseudo-random speckle point spread function generated according to the simple model of the atmosphere described in Section 2.7. The algorithm which effects this is pictured in Figure 4.1. A pseudo-random phase function $\phi_m(\xi)$ is generated (as described below) representing the phase perturbations across the telescope pupil imparted by the atmosphere to incoming light. A unit-magnitude phasor $\alpha^{j\phi_m(\xi)}$ is formed representing the aperture plane field from a point source of unit strength at the origin of object space on the celestial sphere. The unit-magnitude phasor is multiplied by the telescope pupil function $Q(\xi)$ (ϕ1.6.5) to form a coherent transfer function

$$H_{cm}(u) = Q(\bar{\lambda}\ell u)\, e^{j\phi_m(\bar{\lambda}\ell u)} \tag{4.5a}$$

for the atmosphere and telescope combined. The telescope is modelled as being diffraction-limited (§1.6.2), so that

$$Q(\bar{\lambda}\ell u) = P(\bar{\lambda}\ell u) = \text{rect}\left(\frac{\bar{\lambda}\ell u}{D}\right) \tag{4.5b}$$

where D is the diameter of the telescope aperture. The coherent

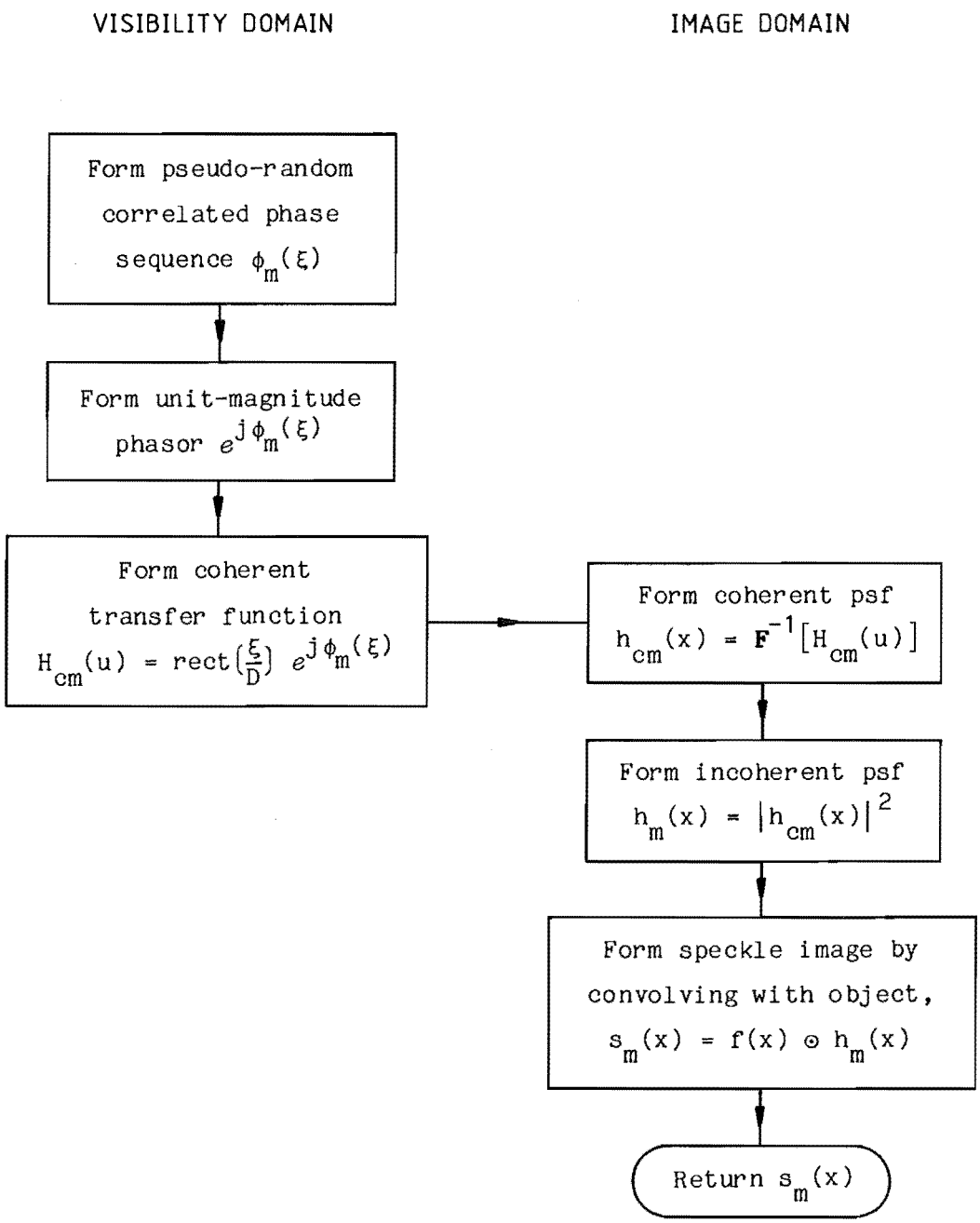


Figure 4.1: Algorithm for simulating isoplanatic, non-photon-limited speckle images.

transfer function is inverse Fourier transformed to give a coherent point spread function $h_{cm}(x)$. The intensity, or squared magnitude, of $h_{cm}(x)$ gives the simulated speckle point spread function

$$h_m(x) = |h_{cm}(x)|^2$$
$$= |F^{-1}[\mathrm{rect}(\tfrac{\bar{\lambda}\ell u}{D})\, e^{j\phi_m(\bar{\lambda}\ell u)}]|^2 . \qquad (4.6)$$

The speckle psf is convolved with the object f(x) to give the isoplanatic, non-photon-limited speckle image

$$s_m(x) = f(x) \odot h_m(x). \qquad (4.7)$$

The convolution is performed in the image domain by directly evaluating the convolution summation

$$s_m(x_p) = \sum_{n=1}^{SIZE} f(\alpha_n)\, h_m(x_p + \ell\alpha_n) \qquad (4.8)$$

rather than through the visibility domain using the FFT. This is done to improve computational efficiency for objects with few nonzero pixels, e.g. binary objects (ϕ5.2), to allow the convolution to be computed non-cyclically [cf. Bracewell 1978 ϕ18], and to allow easy generalisation of the programming to nonisoplanatic speckle images. Most of the computational effort required to generate a single $s_m(x)$ goes into the one-dimensional inverse FFT of $H_{cm}(u)$ which produces $h_{cm}(x)$ (cf. Figure 4.1), unless f(x) has many nonzero pixels (e.g. 20 or more for SIZE = 256).

The pseudo-random phase function $\phi_m(\xi)$ represents the phase perturbations caused by the m^{th} realisation of the turbulent atmosphere. These perturbations are partly correlated in the pupil plane of the telescope, leading to a seeing disc of finite diameter (ϕ2.4). Successive pixels of the one-dimensional $\phi_m(\xi)$ are generated by means of a first-order Markov process, as employed by Robinson [1982 ϕ4.6] in his computer simulations of ultrasonic speckle imaging. Each pixel is formed from a weighted sum of the previous pixel's value and an uncorrelated pseudo-random number β using the formula

$$\phi_m(\xi_1) = \beta_{m1} \,, \tag{4.9a}$$

$$\phi_m(\xi_n) = \gamma\, \phi_m(\xi_{n-1}) + \sqrt{1-\gamma^2}\, \beta_{mn} \,, \quad n > 1. \tag{4.9b}$$

In this formulation the variance of $\phi_m(\xi)$ is equal to the variance of β_m and the correlation length of $\phi_m(\xi)$ is determined by γ. The computer program accepts parameters PR (phase range) and CL (correlation length) and computes γ and δ using

$$\beta = 2\,PR\,(0.5 - r), \tag{4.10a}$$

$$\gamma = 0.5^{1/CL} \tag{4.10b}$$

where r is the pseudo-random number returned by the FORTRAN RAN function [DEC 1984 §D.4.7] with a uniform pdf in the range [0;1). Equation (4.10a) produces a pseudo-random number uniformly distributed between -PR and PR. Equation (4.10b) sets the value of γ so that the correlation between pixels of $\phi_m(\xi)$ separated by CL equals one half. The parameters PR and CL are chosen to give a realistic-looking seeing disc of the desired diameter. The published measurements of atmospheric phase perturbations mentioned in Section 2.2 indicate that seeing PR equal to a few times 2π radians is representative of the Earth's atmosphere. Typical values for PR and CL are quoted in Section 4.7.

The correlation function of the simulated $\phi_m(\xi)$ is negative-exponential because $\phi_m(\xi)$ is generated by a first-order Markov process (4.9). Consequently the atmospheric transfer function

$$H_A(u) = \langle \hat{\mathbf{A}}[e^{j\phi_m(\bar{\lambda}\ell u)}] \rangle \tag{4.11}$$

(cf. 2.8e) of the simulated atmosphere has the negative-exponential form $e^{-k|u|}$ (k is an arbitrary constant), unlike that of the Earth's atmosphere which theoretically has the form $e^{-k|u|^{\frac{5}{3}}}$ (2.11; cf. 2.8e,f). This affects the shape of the simulated seeing disc. Figure 4.2a compares the atmospheric transfer function (4.11) obtained from 5000 speckle images simulated according to (4.9) with the theoretical transfer function of the Earth's atmosphere (2.11) and a Gaussian transfer function, each scaled to identical values of r_o (2.9). These

functions have the forms $e^{-k_1|u|}$, $e^{-k_2|u|^{\frac{5}{3}}}$, $e^{-k_3|u|^2}$ respectively. The corresponding one-dimensional seeing discs, which are the one-dimensional Fourier transforms of the curves in Figure 4.2a, are shown in Figure 4.2b. The simulated seeing disc has unrealistically large tails. However this is not of great importance in the simulations because astronomical speckle processing techniques are in general insensitive to the detailed characteristics of the simulated atmosphere [Labeyrie 1976 ϕ1.2, Bates 1982b ϕ5.2].

Figure 4.3 shows frequency histograms of pixel values in simulated speckle images of a point object (i.e. one occupying only a single pixel of object space). Figure 4.3a was obtained with the simulated telescope aperture D equal to SIZE (cf. ϕ4.7), so that the speckle size was effectively equal to the pixel size and adjacent pixels of $s_m(x)$ were statistically independent of each other. Figure 4.3b was obtained with D = SIZE / 4 so that each speckle covered several pixels. Both of these frequency histograms closely match the theoretical negative-exponential probability density function of quasi-monochromatic speckle images (2.15).

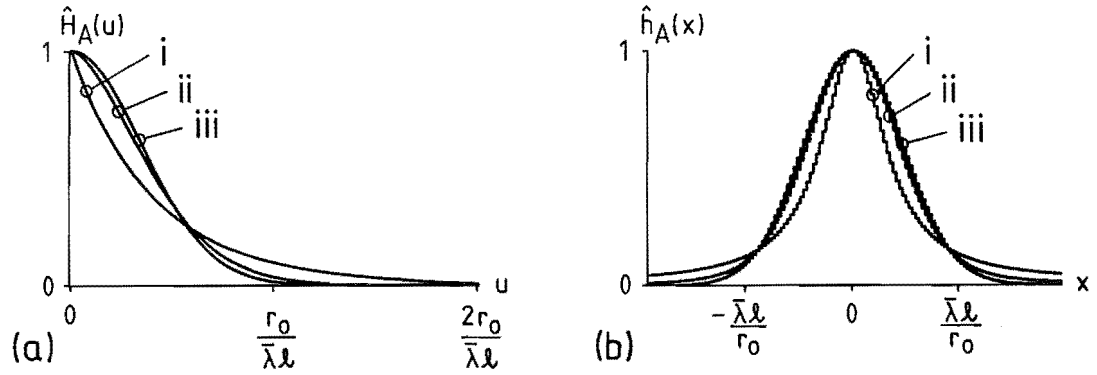


Figure 4.2: Atmospheric transfer functions (a) and seeing discs (b), normalised to identical values of r_0, of

(i) the simulated one-dimensional speckle images, as measured from 5000 speckle images simulated as described in Section 4.4,

(ii) the Earth's atmosphere according to the theoretical expression (2.11),

(iii) a Gaussian transfer function.

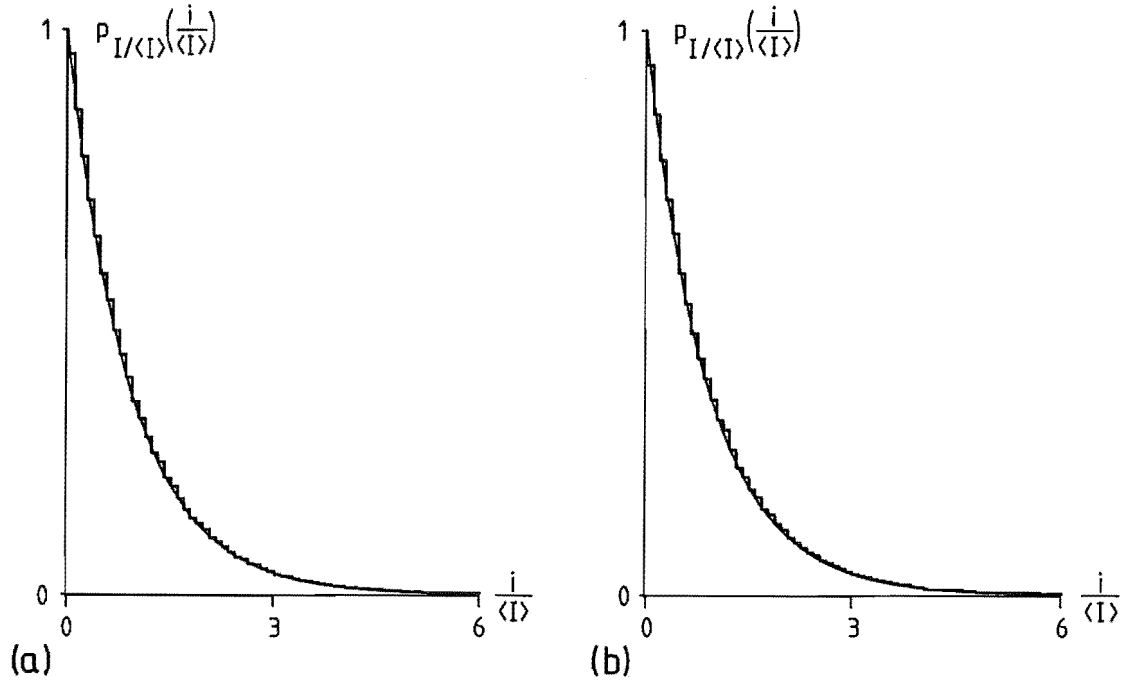


Figure 4.3: Frequency histograms of speckle image intensity $i = s_m(x)$ compiled from 1000 simulated speckle images (φ4.4) of a point object:
(a) with each speckle covering one pixel,
(b) with each speckle covering several pixels.
The theoretical negative-exponential probability density function (2.15c) is drawn as a thin, smoothly curved line in (a) and (b).

4.5 SIMULATING NONISOPLANATISM

Nonisoplanatic speckle images are simulated by modelling a single phase screen elevated to a height h above the telescope aperture (φ2.7). The pseudo-random phase function $\phi_m(\xi)$ is generated as described in Section 4.4, but in the nonisoplanatic case it takes the argument $\xi + h\alpha$ instead of ξ (where α denotes position in the object plane, cf. φ4.3). The speckle point spread function thereby becomes spatially variant and is given by

$$h_m(x,\alpha) = \left| F^{-1}\left[\mathrm{rect}\left(\frac{\bar{\lambda}\ell u}{D}\right) e^{j\phi_m(\bar{\lambda}\ell u + h\alpha)} \right] \right|^2 . \qquad (4.12)$$

A separate $h_m(x,\alpha)$ is generated for each nonzero pixel of the object $f(\alpha)$ (see Figure 4.4). The convolution summation (4.8) is generalised to give the nonisoplanatic speckle image

$$s_m(x_p) = \sum_{n=1}^{SIZE} f(\alpha_n)\, h_m(x_p + \ell\alpha_n, \alpha_n). \qquad (4.13)$$

Two versions of $h_m(x,\alpha)$ corresponding to two pixels in object space separated by an angular distance θ include segments of

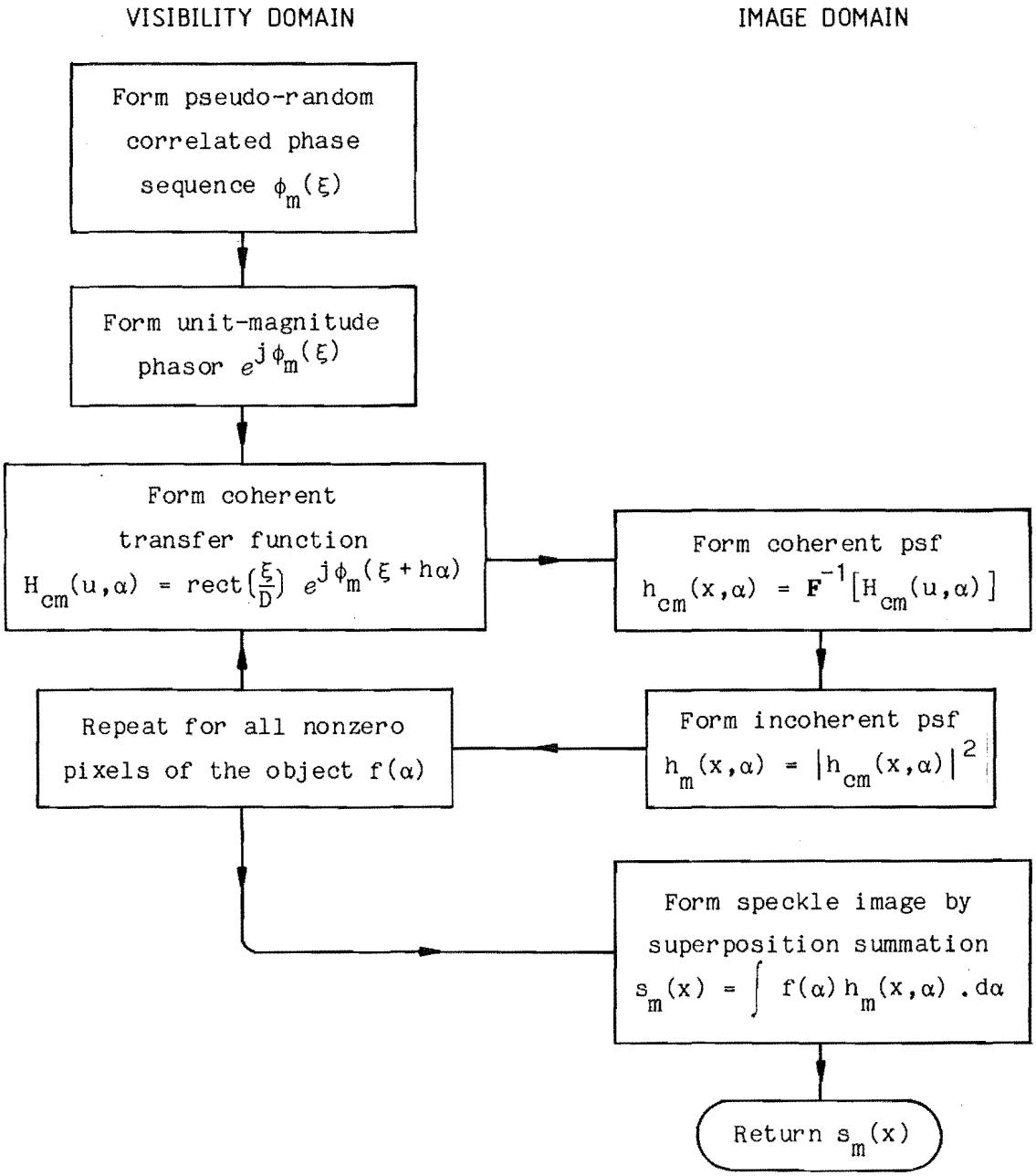


Figure 4.4: Algorithm for simulating nonisoplanatic speckle images.

$\phi_m(\xi)$ displaced by $h\theta$ with respect to each other. When $h = 0$, the segments are identical for all θ and the speckle image (4.13) is isoplanatic. Equation (4.13) then reduces to (4.8). When $h\theta > D$ the segments of $\phi_m(\xi)$ contributing to the two $h_m(x,\alpha)$ are non-overlapping, and hence are uncorrelated with respect to each other if the finite correlation length of $\phi_m(\xi)$ is neglected (cf. $D >> r_o$). The speckle image is then completely nonisoplanatic between the two pixels of object space. The proportional overlap of the segments of $\phi_m(\xi)$ is equivalent to the area ratio $R_a(\theta)$ defined in Section 2.7 (2.36), which for a one-dimensional aperture and a single phase screen at height $h = h$ is given by

$$R_a(\theta) = \begin{cases} \dfrac{D - |h\theta|}{D} & \text{for } |h\theta| \leq D \\ 0 & \text{otherwise.} \end{cases} \tag{4.14}$$

4.6 SIMULATING PHOTON LIMITING

Photon-limited speckle images are simulated according to the semi-classical model of photodetection [Goodman and Belsher 1976, cf. Dainty and Shaw 1974 φ1.1], in which the number n of photons detected for each pixel is a Poisson-distributed random variate with the probability density function

$$p_n(n) = \frac{\bar{n}^n e^{-\bar{n}}}{n!} \tag{4.15a}$$

[cf. Abramowitz and Stegun 1973 φ26.1.22]. The rate $\bar{n}$ is proportional to the classical intensity of the speckle image, i.e.

$$\bar{n}(x) \propto s_m(x). \tag{4.15b}$$

There are N_p photons detected in total in the speckle image. N_p is a Poisson-distributed random variate with mean $\bar{N}_p$ and variance $\bar{N}_p$.

The algorithm used to simulate this process is depicted in Figure 4.5. A non-photon-limited speckle image $s_m(x)$ is generated as described in Sections 4.4 or 4.5. This is treated as the classical

intensity for a photon-limited speckle image $s'_m(x)$, with the probability of each "photon" of $s'_m(x)$ landing in the k^{th} pixel of $s'_m(x)$ made proportional to the value of the k^{th} pixel of $s_m(x)$. The sum of the probabilities over all pixels of $s'_m(x)$ is equal to unity. The mean value $\bar{N}_p$ of N_p determines the average number of photons per speckle image in the simulations. Figure 4.6 shows a typical $s_m(x)$ and three $s'_m(x)$ generated from it with different values of $\bar{N}_p$.

The Poisson-distributed number N_p is simulated in the computer by generating a pseudo-random floating-point number R with an approximately normal probability density function of mean $\bar{N}_p$ and variance $\bar{N}_p$ [cf. Hamming 1973 ∅8.6, Abramowitz and Stegun 1973 ∅26.8.6(2)] and assigning to N_p an integer value according to the rules

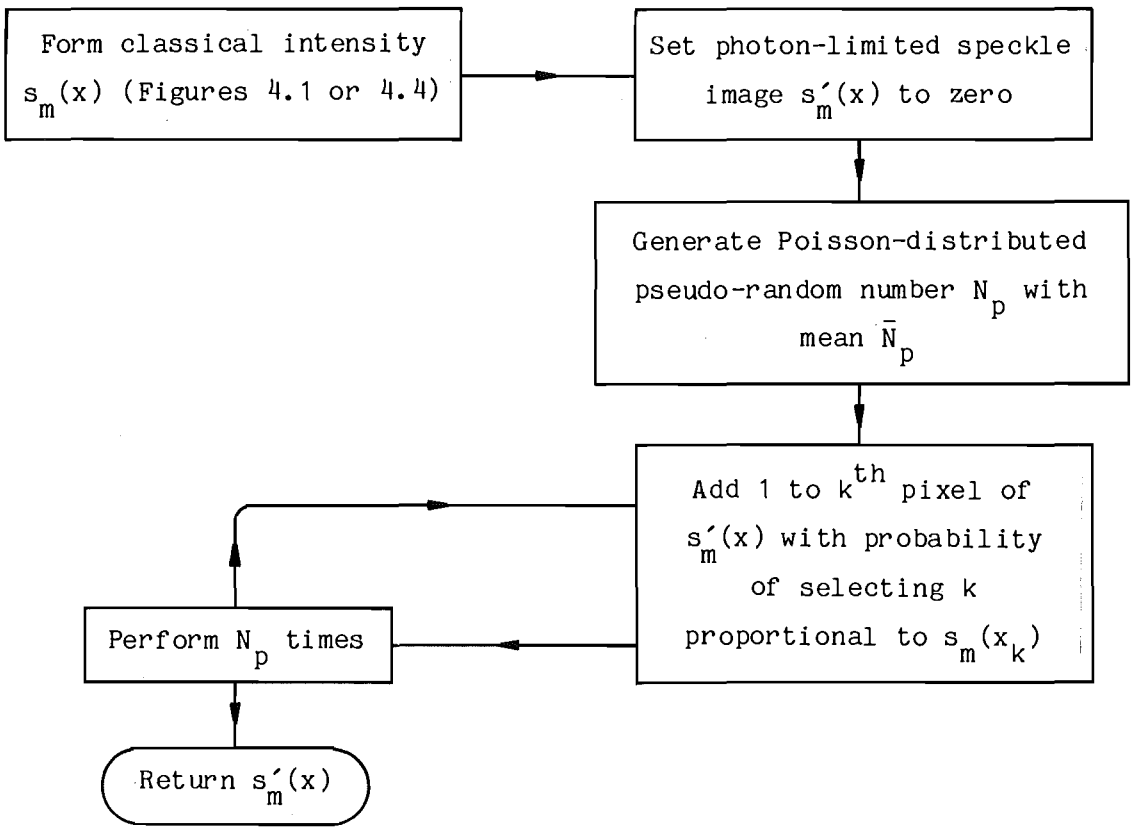


Figure 4.5: Algorithm for simulating photon-limited speckle images with an average of $\bar{N}_p$ photons per speckle image.

$$
\begin{aligned}
R < -0.5 \quad &\Rightarrow \quad N_p \leftarrow 1, \\
-0.5 < R < 0 \quad &\Rightarrow \quad N_p \leftarrow 0, \\
R > 0 \quad &\Rightarrow \quad N_p \leftarrow \text{trunc}(R),
\end{aligned} \tag{4.16}
$$

where trunc(R) denotes truncation of the fractional part from R to leave its integer part. This ad hoc scheme produces pseudo-random N_p whose pdf approximates the Poisson law (4.15a) moderately well for $\bar{N}_p > 1$ and increasingly well as $\bar{N}_p$ increases. It executes equally quickly for any $\bar{N}_p$ (cf. Knuth's [1969 Vol.2 p.117] Algorithm Q which is elegant but computationally expensive for large $\bar{N}_p$).

When $\bar{N}_p$ is small, the randomness or noise in each speckle image originating from photon limiting is greater than that

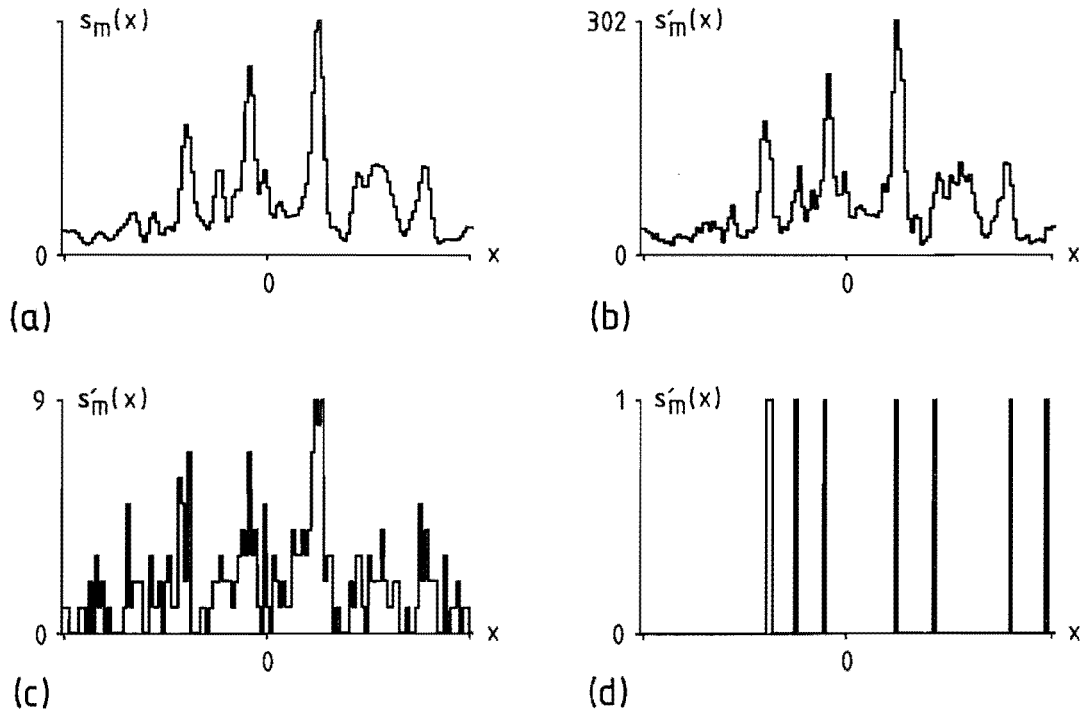


Figure 4.6: Examples of simulated one-dimensional photon-limited speckle images.

(a) Classical intensity: a non-photon-limited speckle image.

(b) - (d) Photon-limited speckle images generated from (a) with $\bar{N}_p$ = 10,000 (b), 300 (c), and 10 (d) photons.

originating from the randomness of the classical intensity $s_m(x)$ (cf. Figure 4.6, (3.42)). Also, when $\bar{N}_p$ is small, most of the computer time required to generate a simulated photon-limited speckle image is taken up generating the classical intensity $s_m(x)$ since this involves at least one Fourier transform (cf. Figures 4.1, 4.4). It is then sensible to generate comparatively few $s_m(x)$ and from these generate as many $s'_m(x)$ as desired, so as to save computer time when $\bar{N}_p$ is small and M (the total number of $s'_m(x)$) is large. For the simulations of Section 5.4, 1000 $s_m(x)$ were stored in a disc file. One $s_m(x)$ was read from the file for each $s'_m(x)$, and the file was rewound and reread as many times as was necessary to generate M $s'_m(x)$.

4.7 SUMMARY OF SIMULATION PARAMETERS

The parameters which control the generation of the computer simulated speckle images are set out here for reference. The table includes typical parameter values used in the simulations reported in Chapters 5 and 6.

Name	Typical Value	Units	Quantity
SIZE	256	pixels	Size of the image domain and visibility domain arrays (ϕ4.3)
PR	15	radians	Phase Range; sets the variance of the pseudo-random phase $\phi_m(\xi)$ (ϕ4.4)
CL	100	pixels	Correlation Length of the pseudo-random phase $\phi_m(\xi)$ (ϕ4.4)
D	64	pixels	Diameter of the simulated telescope aperture (ϕ4.4)
h	0 -	pixels/ pixel	Height of the simulated atmospheric phase screen (ϕ4.5) in units of pixels in ξ-space per pixel in α-space, i.e. pixels in u-space per pixel in x-space by appealing to (4.3, 4.4).
$\bar{N}_p$	5 - 10,000		Average number of photons per simulated photon-limited speckle image (ϕ4.6)
M	300 - 100,000		Number of speckle images averaged in speckle processing

CHAPTER 5

NEW SHIFT-AND-ADD RESULTS

5.1 INTRODUCTION

This chapter presents new results on the performance of simple shift-and-add (ϕ3.6.2) in the context of astronomical speckle imaging. The results are supported by one-dimensional computer simulations and by the processing of two-dimensional astronomical data. Section 5.2 establishes nomenclature and notation for describing the shift-and-add image of a particular class of objects which is important throughout the chapter. Section 5.3 examines the performance of simple shift-and-add under partially isoplanatic conditions. Section 5.4 investigates the effects of photon limiting on the simple shift-and-add image. Conclusions from Sections 5.3 and 5.4 are presented at the end of the respective sections.

The term "shift-and-add" is used to refer specifically to simple shift-and-add, as defined by (3.135) in Section 3.6.2, throughout the rest of this chapter.

5.2 NOMENCLATURE FOR BINARY OBJECTS

In this chapter a binary object is taken to be one which is nonzero in two separated regions of object space (see Figure 5.1a). The object has two components or parts which are sufficiently well separated to be resolved from each other by the imaging instrument under diffraction-limited conditions (cf. ϕ1.6.6). The width of each of the parts is substantially less than the separation between them. The parts themselves may individually be resolved or unresolved. The binary object with both parts unresolved, as illustrated in Figures 5.1a and b, is a useful test object for studies of speckle processing because it is a simple object containing non-trivial detail and is representative of an important class of astronomical objects, namely binary stars [cf. McAlister and Hartkopf 1984].

A binary object with unresolved parts is adequately represented by the sum of two delta functions, i.e.

$$f(\underset{\sim}{\alpha}) = \delta(\underset{\sim}{\alpha} - \underset{\sim}{\alpha}_0) + a\,\delta(\underset{\sim}{\alpha} - \underset{\sim}{\alpha}_0 - \underset{\sim}{\theta}) \tag{5.1a}$$

in object space (see Figure 5.1a) or

$$f(\underset{\sim}{x}) = \delta(\underset{\sim}{x} - \underset{\sim}{x}_0) + a\,\delta(\underset{\sim}{x} - \underset{\sim}{x}_0 - \underset{\sim}{b}) \tag{5.1b}$$

in image space (cf. ϕ2.2) (see Figure 5.1b), where

$$\underset{\sim}{x} = -\ell\underset{\sim}{\alpha} \tag{5.2}$$

(cf. 2.1) relates positions in the object and image spaces. The object is characterised by the two parameters a and $\underset{\sim}{\theta}$ in object space, or equivalently by a and $\underset{\sim}{b}$ in image space, where

$$\underset{\sim}{b} = -\ell\underset{\sim}{\theta}. \tag{5.3}$$

The primary peak of the object is the unit-strength impulse at $\underset{\sim}{\alpha} = \underset{\sim}{\alpha}_0$ in object space, and the secondary peak of the object is the impulse at $\underset{\sim}{\alpha} = \underset{\sim}{\alpha}_0 + \underset{\sim}{\theta}$ with strength a obeying

$$0 \leq a \leq 1. \tag{5.4}$$

$\underset{\sim}{\theta}$ or $\underset{\sim}{b}$ represents the separation between the two parts. The absolute location $\underset{\sim}{\alpha}_0$ of the object is arbitrary when considering shift-and-add because changes in its value do not affect the shift-and-add image.

The shift-and-add image of a binary object possesses four components (see Figure 5.1c): the primary peak, the secondary peak, the ghost peak, and the fog (cf. ϕ3.6.2). The primary peak is centred about $\underset{\sim}{x} = \underset{\sim}{0}$ in image space and the secondary and ghost peaks are centred about $\underset{\sim}{x} = \underset{\sim}{b}$ and $\underset{\sim}{x} = -\underset{\sim}{b}$ respectively. The shift-and-add image may be characterised by parameters y_p, y_s, y_g, and y_f describing the heights of the primary peak above the fog, the secondary peak above the fog, the ghost peak above the fog, and the fog at the primary peak, respectively, as shown in Figure 5.1c. Normalised versions of these parameters are defined with respect to the height of

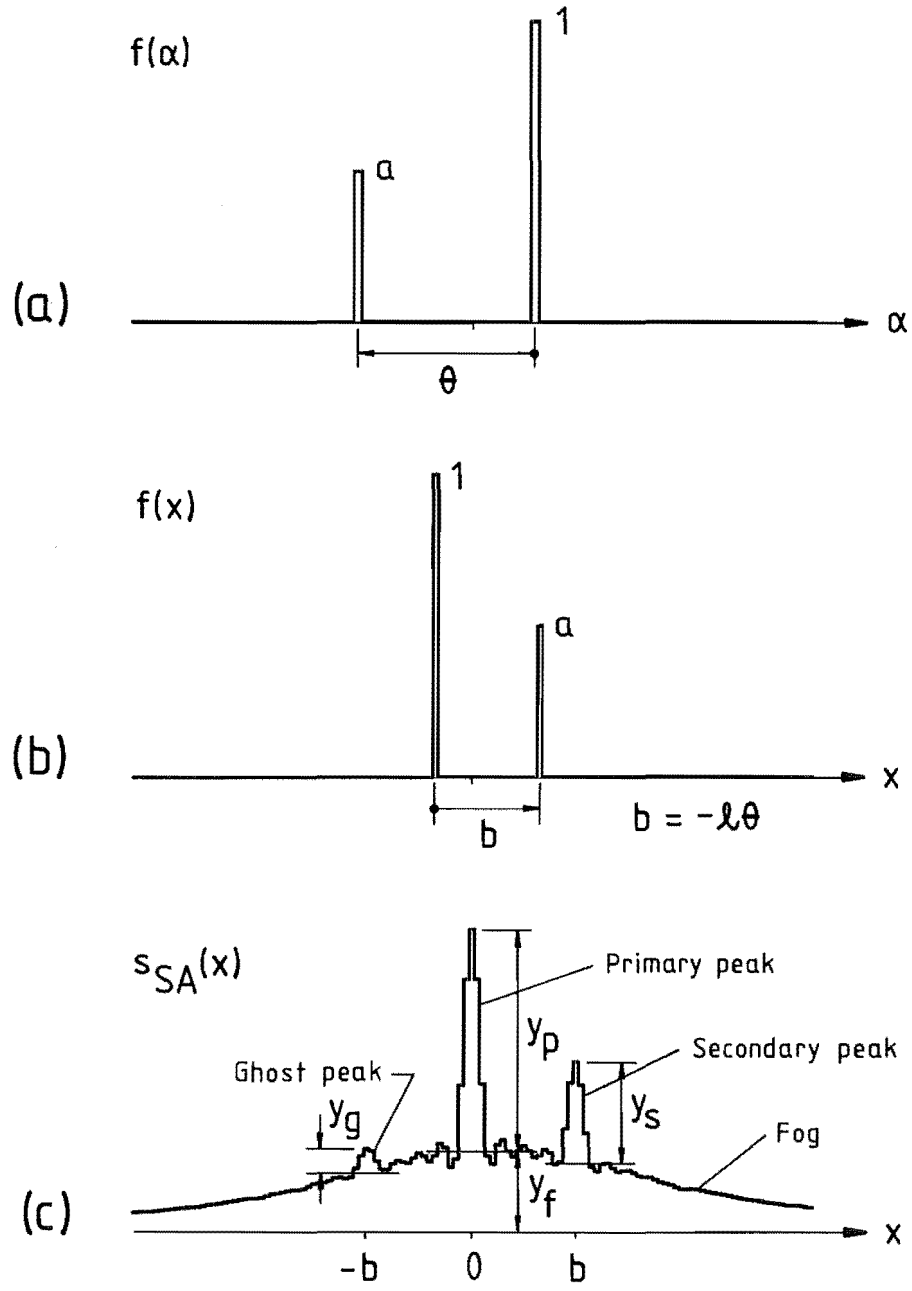


Figure 5.1: A binary object represented in object space (a) and in image space (b), and its shift-and-add image (c), showing the components of the object and the image and the parameters describing the components.

the primary peak above the fog by

$$\hat{y}_p = \frac{y_p}{y_p} = 1, \tag{5.5a}$$

$$\hat{y}_s = \frac{y_s}{y_p}, \tag{5.5b}$$

$$\hat{y}_g = \frac{y_g}{y_p}, \tag{5.5c}$$

$$\hat{y}_f = \frac{y_f}{y_p}. \tag{5.5d}$$

y_p, y_s, and y_g represent the heights of the peaks in the defogged shift-and-add image (ϕ3.6.2). The defogging can be performed by deconvolution or by subtraction with essentially equivalent results, provided the fog is accurately estimated (cf. Figure 3.10d). The subtractive approach is implied in the definitions of y_p, y_s, and y_g above (cf. Figure 5.1c). This approach is convenient for a binary object with unresolved parts, since the fog levels in $s_{SA}(\underset{\sim}{x})$ can be readily estimated by averaging the pixel values through several of the sidelobes on either side of the base of each peak in $s_{SA}(\underset{\sim}{x})$ (Figure 5.1c).

5.3 SHIFT-AND-ADD WITH PARTIALLY ISOPLANATIC SPECKLE IMAGES

Shift-and-add in its basic form (ϕ3.6.2) requires an object to have a dominatingly bright, unresolved reference part for it to be imaged by shift-and-add without being significantly contaminated by ghosting. In astronomical speckle imaging, any unresolved star that is sufficiently bright relative to its surroundings can serve as a reference source for shift-and-add. The object comprising the unresolved star and its immediate surroundings can thereby be faithfully imaged by shift-and-add. The role of the reference star is similar to its role in speckle holography (ϕ3.4.1). However shift-and-add is computationally much simpler than speckle holography, and is therefore to be preferred except in situations where the number of photons detected per speckle image is so low that shift-and-add is ineffective (cf. ϕ5.4).

As with speckle holography, the extent of the region about the reference star which can be imaged to high resolution by shift-and-add is governed by the size of the isoplanatic patch, which is itself determined by the state of the atmospheric seeing. This section investigates the effects of nonisoplanatism on shift-and-add. Section 5.3.1 examines one-dimensional shift-and-add images formed from partially isoplanatic computer simulated speckle images. Section 5.3.2 presents a simple mathematical model of partially isoplanatic speckle images, shows the consequences for the shift-and-add image under the model, and compares these consequences with the simulations of Section 5.3.1. In Section 5.3.3 the loss of resolution due to nonisoplanatism in the simulations of Section 5.3.1 is further examined. Section 5.3.4 describes a modification of shift-and-add which for a particular type of object relaxes the requirement that the reference part of the object be dominatingly bright, and presents a computational example illustrating how the modified shift-and-add algorithm might be applied in astronomical practice. In Section 5.3.5 shift-and-add is applied to partially isoplanatic two-dimensional speckle images of the double star α Scorpii, and the results are discussed. Conclusions on the effects of nonisoplanatism on shift-and-add follow in Section 5.3.6.

5.3.1 COMPUTER SIMULATIONS WITH A BINARY OBJECT

Figure 5.2 depicts a series of one-dimensional shift-and-add images formed according to (3.135) from partially isoplanatic computer simulated speckle images of the binary object shown in Figure 5.2a. This object is faithfully imaged by shift-and-add in isoplanatic conditions since its unresolved primary peak is considerably brighter than its secondary peak. The secondary peak of the object is also unresolved in this example, allowing the effects of partial isoplanatism to be readily interpreted in the shift-and-add images shown in Figures 5.2b-f. The computer simulated speckle images were generated as described in Section 4.5 from the physical model of the atmosphere presented in Section 2.7. The degree of isoplanatism between the two parts of the object varies from 1 (completely isoplanatic) to 0 (completely nonisoplanatic) in Figures 5.2b-f. The quantity R_a quoted in the figures is the area ratio (4.14) controlling

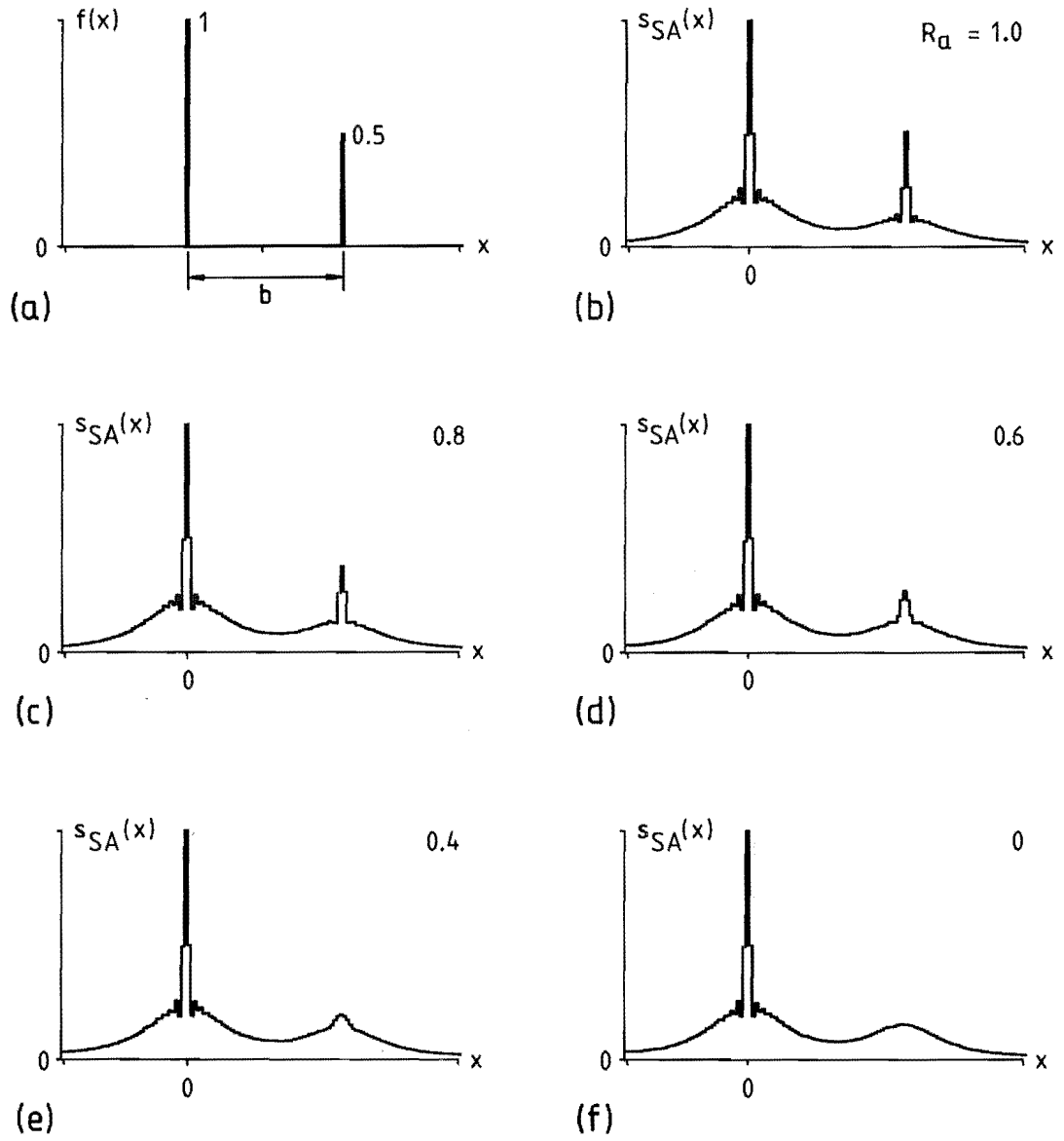


Figure 5.2: Shift-and-add with partially isoplanatic one-dimensional simulated speckle images.

(a) Object.

(b) - (f) Shift-and-add images, each formed from 4000 speckle images of (a), with R_a = 1.0 (b), 0.8 (c), 0.6 (d), 0.4 (e), and 0 (f).

The speckle images were generated with SIZE = 256, PR = 10, CL = 100, D = SIZE/2, h as defined by R_a, θ = b = 50 pixels (see Chapter 4). SIZE/ 2 pixels of image space are displayed.

the degree of isoplanatism in the simulations.

Nonisoplanatism has no discernible effect on the primary peak of the shift-and-add images (Figures 5.2b-f). The secondary peak of the shift-and-add images diminishes in height and disappears into the fog as the degree of isoplanatism between the two parts of the object is reduced from 1 to 0. There is a slight broadening of the secondary peak as the degree of isoplanatism is reduced. The variation in height of the secondary peak above the fog, relative to the height of the primary peak above the fog (cf. ϕ5.2), is plotted as a function of R_a in Figure 5.3. The variation agrees well with the relation

$$\hat{y}_s = R_a^2 \cdot \hat{y}_s|_{R_a=0} \tag{5.6}$$

derived theoretically by Bagnuolo [1984] and shown as the dotted line in Figure 5.3. Bagnuolo [1984] shows that the normalised height ratio

$$R_h = \hat{y}_s \; / \; \hat{y}_s|_{R_a=0} \tag{5.7}$$

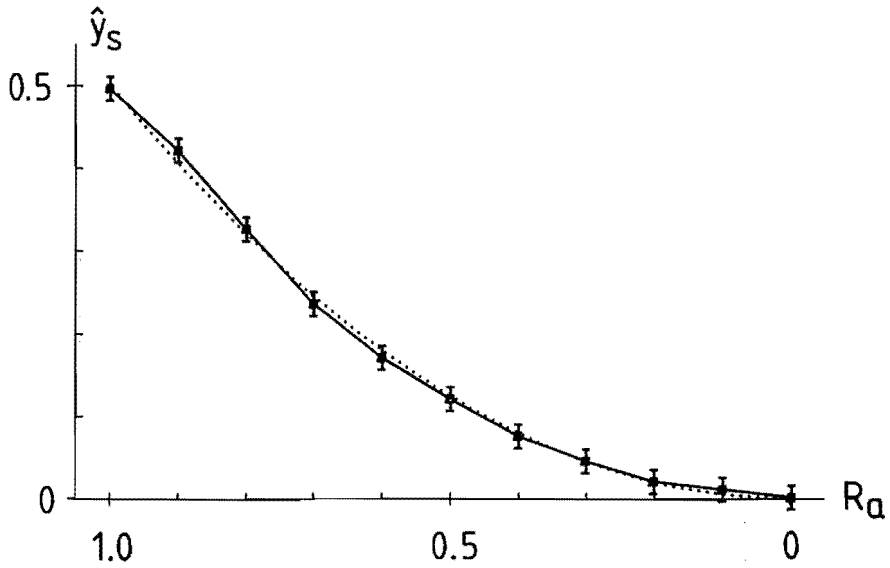


Figure 5.3: Variation of secondary peak height with degree of isoplanatism in the shift-and-add images of Figure 5.2. The dotted line shows the theoretical prediction of Bagnuolo [1984].

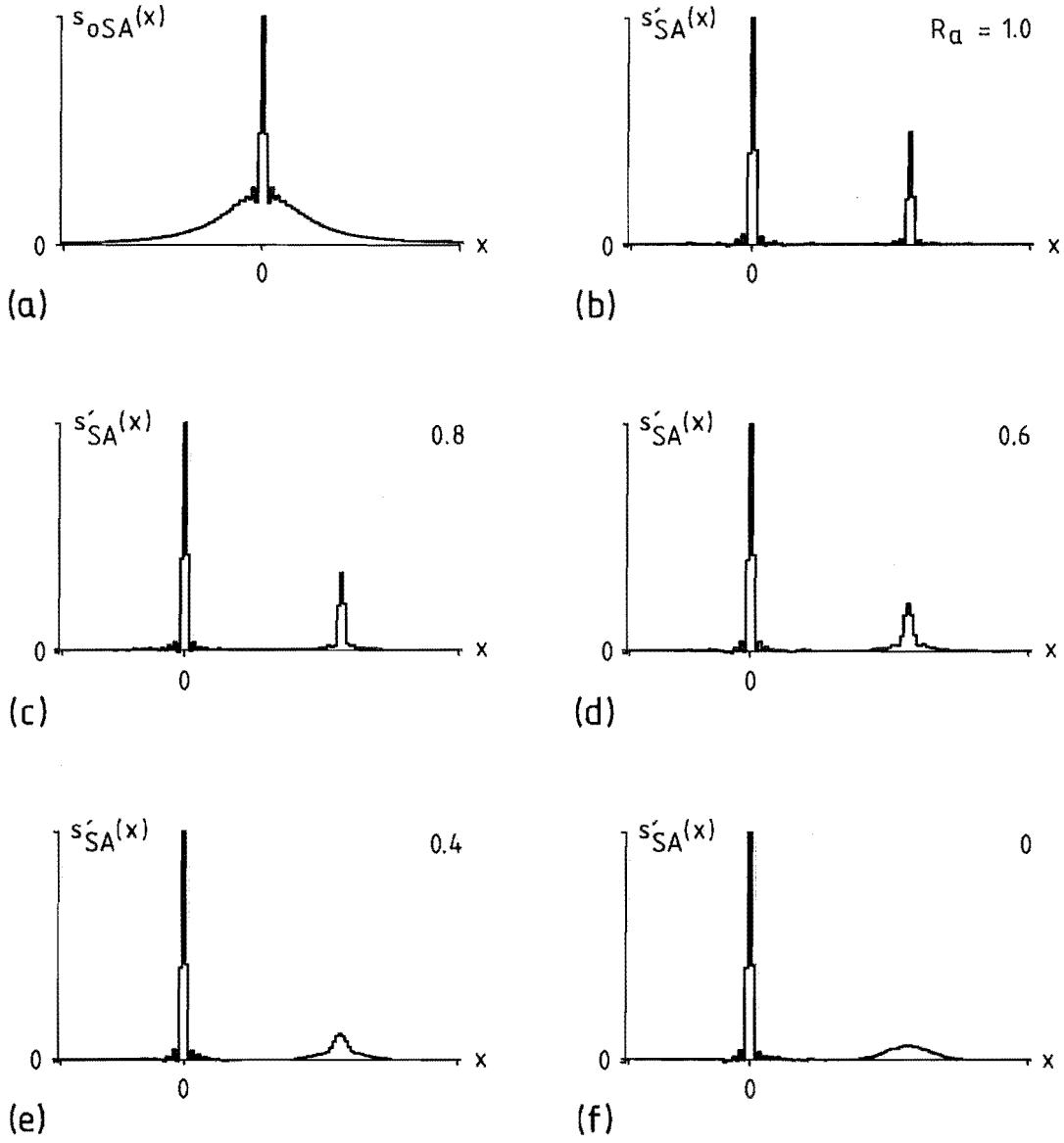


Figure 5.4: Defogged versions of the shift-and-add images of Figure 5.2.

(a) Shift-and-add image formed from 4000 simulated speckle images of a point object.

(b) - (f) Defogged shift-and-add images formed by deconvolving (a) from Figures 5.2b-f respectively, by Wiener filtering, then convolving the result by the diffraction-limited point spread function of the simulated telescope.

The speckle images for (a) were generated as described in the caption to Figure 5.2.

is equivalent to the height ratio $R_h(\underset{\sim}{\theta})$ (2.34) for speckle interferometry defined in Section 2.6, which for atmospheric turbulence concentrated in a single thin layer is given by (2.38) (cf. 5.6). The height ratio is a useful quantitative measure of the degree of isoplanatism of the speckle images with respect to the spatial interval separating the two parts of the object. The interval at which R_h drops to effectively zero defines the size of the isoplanatic patch for shift-and-add.

The background fog in a shift-and-add image $s_{SA}(\underset{\sim}{x})$ which results from the low spatial frequency excess in $h_{SA}(\underset{\sim}{x})$ can be suppressed by deconvolving from $s_{SA}(\underset{\sim}{x})$ the shift-and-add image $s_{oSA}(\underset{\sim}{x})$ of an unresolved reference object viewed through statistically similar seeing (cf. 3.140). Figure 5.4 shows the shift-and-add images of Figure 5.2b-f defogged according to (3.140), using Figure 5.4a as $s_{oSA}(\underset{\sim}{x})$. The absence of the fog in Figures 5.4b-f allows another effect of partial isoplanatism on shift-and-add to be seen. As the degree of isoplanatism is reduced and the secondary peak diminishes in height, a broad, low hump of residual fog appears about the secondary peak of the shift-and-add images. No residual fog appears about the primary peak.

5.3.2 A SIMPLE MATHEMATICAL MODEL OF PARTIAL ISOPLANATISM

The physical model of the atmosphere used to generate nonisoplanatic speckle images in the computer simulations of Section 5.3.1 is computationally convenient and conceptually well linked to the mechanisms of optical propagation through the atmosphere. However it does not readily lead to a mathematical description of the shift-and-add image under partially isoplanatic conditions. This subsection presents a mathematical model of partially isoplanatic speckle images, which leads straightforwardly to a mathematical description of the shift-and-add image formed from them. The consequences of the mathematical description are considered and compared with the results of the simulations of Section 5.3.1.

5.3.2.1 Speckle Images

A nonisoplanatic speckle image is formally described (cf. ϕ1.2.1) by a superposition integral over the object with a spatially varying point spread function, expressible as

$$s_m(\underset{\sim}{x}) = \int f(\underset{\sim}{\alpha})\, h_m(\underset{\sim}{x} + \ell\underset{\sim}{\alpha}, \underset{\sim}{\alpha})\, .d\underset{\sim}{\alpha} \qquad (5.8)$$

(cf. 1.8, 2.1, 2.27). In this subsection the object is represented as a sum of parts,

$$f(\underset{\sim}{x}) = \sum_{n=1}^{N} f_n(\underset{\sim}{x}), \qquad (5.9)$$

each of which has an extent that is nonzero but small enough for the speckle point spread function to be essentially invariant across it. The speckle image (5.8) can then be written as

$$s_m(\underset{\sim}{x}) = \sum_{n=1}^{N} f_n(\underset{\sim}{x}) \circledcirc h_{mn}(\underset{\sim}{x}), \qquad (5.10)$$

where $h_{mn}(\underset{\sim}{x})$ is the speckle psf associated with the contribution of the n^{th} part of the object to the m^{th} speckle image. Equations (5.9) and (5.10) can in principle be applied to any object by arbitrarily dividing it into parts, but they are particularly appropriate to binary objects as defined in Section 5.2. Each $h_{mn}(\underset{\sim}{x})$ is modelled as a linear combination of the speckle psf $h_{m1}(\underset{\sim}{x})$ corresponding to a particular reference part $f_1(\underset{\sim}{x})$ of the object, and a $g_{mn1}(\underset{\sim}{x})$ which is statistically independent of $h_{m1}(\underset{\sim}{x})$. $g_{mn1}(\underset{\sim}{x})$ represents the change in form of the speckle psf caused by $f_1(\underset{\sim}{x})$ and $f_n(\underset{\sim}{x})$ being viewed through different parts of the atmosphere. In fact, $g_{mn1}(\underset{\sim}{x})$ characterises the differences, as regards the form of the speckle psf, between these two parts of the atmosphere (cf. the singly cross-hatched regions of Figure 2.5). The model is expressed symbolically as

$$h_{mn}(\underset{\sim}{x}) = \beta_{mn1}\, h_{m1}(\underset{\sim}{x}) + (1 - \gamma_{mn1})\, g_{mn1}(\underset{\sim}{x}) \qquad (5.11)$$

where β_{mn1} and γ_{mn1} are random numbers in the range [0;1] obeying the relations

$$\langle \beta_{mnl} \rangle = \langle \gamma_{mnl} \rangle = \beta_{nl} = \beta_{ln} , \tag{5.12a}$$

$$\beta_{mnn} = \gamma_{mnn} = 1, \tag{5.12b}$$

and $g_{mnl}(\underset{\sim}{x})$ is a speckle psf from the same statistical ensemble as $h_{ml}(\underset{\sim}{x})$ but statistically independent of it. $\beta_{nl} = \beta_{ln}$ is the <u>degree of isoplanatism</u> between parts l and n of the object. Substituting (5.11) into (5.10) gives

$$s_m(\underset{\sim}{x}) = \sum_{n=1}^{N} f_n(\underset{\sim}{x}) \odot \left[\beta_{mnl} h_{ml}(\underset{\sim}{x}) + (1 - \gamma_{mnl}) g_{mnl}(\underset{\sim}{x}) \right], \tag{5.13}$$

which provides a mathematical model of a partially isoplanatic speckle image.

5.3.2.2 The Shift-and-Add Image

The reference part $f_1(\underset{\sim}{x})$ in the previous subsection is taken to be the reference part of the object for shift-and-add (∅3.6.2). The shift-and-add image is described by

$$s_{SA}(\underset{\sim}{x}) = \langle s_m(\underset{\sim}{x} + \underset{\sim}{x}_m) \rangle_M \tag{5.14}$$

(3.135) where $\underset{\sim}{x}_m$ is the position of the brightest pixel in $s_m(\underset{\sim}{x})$. Substituting (5.13) into (5.14) yields

$$\begin{aligned} s_{SA}(\underset{\sim}{x}) &= \Big\langle \sum_{n=1}^{N} f_n(\underset{\sim}{x}) \odot \left[\beta_{mnl} h_{ml}(\underset{\sim}{x} + \underset{\sim}{x}_m) \right. \\ &\qquad \left. + (1 - \gamma_{mnl}) g_{mnl}(\underset{\sim}{x} + \underset{\sim}{x}_m) \right] \Big\rangle_M \\ &= \sum_{n=1}^{N} f_n(\underset{\sim}{x}) \odot \langle \beta_{mnl} h_{ml}(\underset{\sim}{x} + \underset{\sim}{x}_m) \rangle_M \\ &\qquad + \sum_{n=1}^{N} f_n(\underset{\sim}{x}) \odot \langle (1 - \gamma_{mnl}) g_{mnl}(\underset{\sim}{x} + \underset{\sim}{x}_m) \rangle_M . \end{aligned} \tag{5.15}$$

If $f_1(\underset{\sim}{x})$ is a good reference for shift-and-add, so that $\underset{\sim}{x}_m$ is effectively influenced only by the component of $s_m(\underset{\sim}{x})$ originating from $f_1(\underset{\sim}{x})$, then $\underset{\sim}{x}_m$ is statistically dependent on $h_{ml}(\underset{\sim}{x})$ but is statistically independent of $g_{mnl}(\underset{\sim}{x})$. The averages in the first summation on the right hand side of (5.15) are each therefore similar

to the shift-and-add psf for isoplanatic conditions (3.136) and each contain a diffraction-limited component, whereas the averages in the second summation yield zero for n = 1 and a smoothly varying function with no significant spatial frequencies above the seeing limit for each n ≠ 1. Thus the second summation in (5.15) contributes only to the fog in $s_{SA}(\underset{\sim}{x})$. Writing

$$s_f(\underset{\sim}{x}) = \sum_{n=1}^{N} f_n(\underset{\sim}{x}) \odot \langle (1 - \gamma_{mnl})\, g_{mnl}(\underset{\sim}{x} + \underset{\sim}{x}_m) \rangle_M \tag{5.16}$$

for the fog-like second term of (5.15), and also making the assumption that

$$\langle \beta_{mnl}\, h_{ml}(\underset{\sim}{x} + \underset{\sim}{x}_m) \rangle_M = \langle \beta_{mnl} \rangle_M \langle h_{ml}(\underset{\sim}{x} + \underset{\sim}{x}_m) \rangle_M , \tag{5.17}$$

(5.15) becomes

$$\begin{aligned} s_{SA}(\underset{\sim}{x}) &= \sum_{n=1}^{N} \left(f_n(\underset{\sim}{x}) \odot \langle \beta_{mnl} \rangle_M \langle h_{ml}(\underset{\sim}{x} + \underset{\sim}{x}_m) \rangle_M \right) + s_f(\underset{\sim}{x}) \\ &= \left(\sum_{n=1}^{N} \beta_{nl} f_n(\underset{\sim}{x}) \right) \odot \langle h_{ml}(\underset{\sim}{x} + \underset{\sim}{x}_m) \rangle_M + s_f(\underset{\sim}{x}) \\ &= \left(\sum_{n=1}^{N} \beta_{nl} f_n(\underset{\sim}{x}) \right) \odot h_{SA}(\underset{\sim}{x}) + s_f(\underset{\sim}{x}) \end{aligned} \tag{5.18}$$

where

$$h_{SA}(\underset{\sim}{x}) = \langle h_{ml}(\underset{\sim}{x} + \underset{\sim}{x}_m) \rangle_M \tag{5.19}$$

is the shift-and-add psf (3.136) applying to the reference part of the object.

The background fog in the shift-and-add image which results from the low spatial frequency excess in $h_{SA}(\underset{\sim}{x})$ can be suppressed by deconvolving from $s_{SA}(\underset{\sim}{x})$ the shift-and-add image of an unresolved reference object (∅3.6.2). Substituting (5.18) into the appropriate equation (3.140) gives

$$s'_{SA}(\underset{\sim}{x}) \approx \Big(\sum_{n=1}^{N} \beta_{n1} f_n(\underset{\sim}{x}) \Big) \odot h_{SA}(\underset{\sim}{x}) \odot^{-1} h_{oSA}(\underset{\sim}{x}) \odot h_D(\underset{\sim}{x}) + s_f(\underset{\sim}{x}) \odot^{-1} h_{oSA}(\underset{\sim}{x}) \odot h_D(\underset{\sim}{x})$$

$$\approx \Big(\sum_{n=1}^{N} \beta_{n1} f_n(\underset{\sim}{x}) \Big) \odot h_D(\underset{\sim}{x}) + s'_f(\underset{\sim}{x}), \qquad (5.20)$$

in which a trivial scale factor has been dropped and the simplification

$$h_{SA}(\underset{\sim}{x}) \odot^{-1} h_{oSA}(\underset{\sim}{x}) \approx \delta(\underset{\sim}{x}) \qquad (5.21)$$

has been made. Equation (5.21) is a consequence of the assumption introduced earlier in this subsection that $f_1(\underset{\sim}{x})$ is a good shift-and-add reference, implying that ghosting is negligible.

5.3.2.3 Interpretation and Comparison with the Simulations

Equation (5.18) predicts that the shift-and-add image formed under partially isoplanatic conditions contains detail corresponding to each part of the object, attenuated by a height ratio R_h (cf. 5.7) equal to the degree of isoplanatism β_{n1} existing between each part of the object and the reference part. The detail corresponding to the reference part of the object is not attenuated since $\beta_{11} = 1$ (cf. 5.12b, 5.12a). The remainder of the energy from each part of the object is smeared out into an additional background fog $s_f(\underset{\sim}{x})$. Equation (5.20) predicts that the defogged shift-and-add image resembles the true image, except that parts of the object away from the reference part are attenuated by the corresponding degrees of isoplanatism, plus a residual fog $s'_f(\underset{\sim}{x})$ arising from the nonisoplanatism.

These predictions agree well with the simulated results of Section 5.3.1, except that the mathematical model does not predict the slight broadening of the secondary peak which occurs as the degree of isoplanatism is reduced towards zero. However the broadening is only noticeable in Figures 5.2d-f and 5.3d-f, for which the height ratio R_h (5.7) is less than 0.4. For $R_h > 0.4$ the broadening is insignificant

and the simple mathematical model appears to provide an accurate description of the effects of partial isoplanatism on shift-and-add.

An important conclusion to be drawn from the mathematical model is that no spurious detail is created in the shift-and-add image by partial isoplanatism. Detail is merely attenuated and the background fog is increased slightly by the presence of $s_f(\underset{\sim}{x})$ in (5.18).

5.3.3 LOSS OF RESOLUTION: EMPIRICAL RESULTS

The broadening of the secondary peak of the simulated shift-and-add images shown in Section 5.3.1 implies that partial isoplanatism causes a loss of resolution as well as an attenuation of height of detail in shift-and-add images formed under partially isoplanatic conditions. In this subsection the loss of resolution experienced by shift-and-add is compared empirically with that suffered by speckle interferometry.

5.3.3.1 Roddier et al's [1982] Theory for Speckle Interferometry

Roddier et al [1982] present a comprehensive theoretical treatment of the effects of partial isoplanatism on speckle interferometry. Their development is briefly summarised here (see also ϕ2.6).

The speckle visibility under nonisoplanatic conditions is given by the Fourier transform with respect to $\underset{\sim}{x}$ of (5.8), which on appealing to the shift theorem for the Fourier transform [Bracewell 1978 pp.122,244] is given by (cf. 1.12a, 5.2)

$$S_m(\underset{\sim}{u}) = \ell \int f(\underset{\sim}{\alpha})\, H_m(\underset{\sim}{u},\underset{\sim}{\alpha})\, e^{j2\pi \underset{\sim}{u}.\ell\underset{\sim}{\alpha}} .d\underset{\sim}{\alpha}. \qquad (5.22)$$

Speckle interferometry (3.12) forms the averaged power spectrum

$$S_{LA}(\underset{\sim}{u}) \approx \langle |S_m(\underset{\sim}{u})|^2 \rangle \quad \text{(for large M)}$$

$$= \langle \ell^2 \iint f(\underset{\sim}{\alpha}') f(\underset{\sim}{\alpha}'') H_m^*(\underset{\sim}{u},\underset{\sim}{\alpha}') H_m(\underset{\sim}{u},\underset{\sim}{\alpha}'') \cdot e^{j2\pi \underset{\sim}{u}.\ell(\underset{\sim}{\alpha}'' - \underset{\sim}{\alpha}')} \cdot d\underset{\sim}{\alpha}' \cdot d\underset{\sim}{\alpha}'' \rangle$$

$$= \ell^2 \iint f(\underset{\sim}{\alpha}') f(\underset{\sim}{\alpha} + \underset{\sim}{\alpha}') \langle H_m^*(\underset{\sim}{u},\underset{\sim}{\alpha}') H_m(\underset{\sim}{u},\underset{\sim}{\alpha}' + \underset{\sim}{\alpha}) \rangle \cdot e^{j2\pi \underset{\sim}{u}.\ell\underset{\sim}{\alpha}} \cdot d\underset{\sim}{\alpha}' \cdot d\underset{\sim}{\alpha}$$

where $\underset{\sim}{\alpha} = \underset{\sim}{\alpha}'' - \underset{\sim}{\alpha}'$

$$= \ell^2 \int ff(\underset{\sim}{\alpha}) \langle H_m^*(\underset{\sim}{u},\underset{\sim}{\alpha}') H_m(\underset{\sim}{u},\underset{\sim}{\alpha}' + \underset{\sim}{\alpha}) \rangle e^{j2\pi \underset{\sim}{u}.\ell\underset{\sim}{\alpha}} \cdot d\underset{\sim}{\alpha}, \tag{5.23}$$

where

$$ff(\underset{\sim}{\alpha}) = \mathbf{A}[f(\underset{\sim}{\alpha})] \tag{5.24}$$

is the autocorrelation of the object $f(\underset{\sim}{\alpha})$ and the speckle cross-spectrum $\langle H_m^*(\underset{\sim}{u},\underset{\sim}{\alpha}') H_m(\underset{\sim}{u},\underset{\sim}{\alpha}' + \underset{\sim}{\alpha}) \rangle$ is assumed to be independent of $\underset{\sim}{\alpha}'$. Equation (5.23) is the general expression for $S_{LA}(\underset{\sim}{u})$ in terms of the speckle cross-spectrum [cf. Roddier et al 1982 eq.(4)]. Substituting Roddier et al's [1982] result for the speckle cross-spectrum (2.30) into (5.23) gives

$$S_{LA}(\underset{\sim}{u}) \approx \ell^2 \int ff(\underset{\sim}{\alpha}) \left(|H_{LE}(\underset{\sim}{u})|^2 + \sigma(\underset{\sim}{\alpha})\, Q_4(\bar{\lambda}\ell\underset{\sim}{u}, \bar{h}\underset{\sim}{\alpha}) \right) e^{j2\pi \underset{\sim}{u}.\ell\underset{\sim}{\alpha}} \cdot d\underset{\sim}{\alpha} \tag{5.25}$$

[cf. Roddier et al 1982 eq.(15)]. Equation (5.25) states that, in the speckle interferometry image $s_{LA}(\underset{\sim}{x}) \leftrightarrow S_{LA}(\underset{\sim}{u})$, each point of the object's autocorrelation $ff(\underset{\sim}{\alpha})$ contributes to $s_{LA}(\underset{\sim}{x})$ through the transfer function

$$|H_{LE}(\underset{\sim}{u})|^2 + \sigma(\underset{\sim}{\alpha})\, Q_4(\bar{\lambda}\ell\underset{\sim}{u}, \bar{h}\underset{\sim}{\alpha}). \tag{5.26}$$

Note that

$$Q_4(\bar{\lambda}\ell\underset{\sim}{u},\bar{h}\underset{\sim}{\alpha}) = Q_4(\underset{\sim}{\xi},\bar{h}\underset{\sim}{\alpha}) = \mathbf{A}\left[Q^*(\underset{\sim}{\xi})\,Q(\underset{\sim}{\xi}+\bar{h}\underset{\sim}{\alpha})\right] \tag{5.27}$$

(cf. 3.3a) is the incoherent transfer function (cf. ϕ1.6.4) associated with a pupil function $Q^*(\underset{\sim}{\xi})\,Q(\underset{\sim}{\xi}+\bar{h}\underset{\sim}{\alpha})$. The first term in the transfer function (5.26) extends only to the seeing limit in Fourier space, and is responsible for the background fog in $s_{LA}(\underset{\sim}{x})$. The second term in (5.26), which transmits the detail of $ff(\underset{\sim}{\alpha})$ appearing on top of the fog in $s_{LA}(\underset{\sim}{x})$, has an extent in Fourier space governed by $Q_4(\bar{\lambda}\ell\underset{\sim}{u},\bar{h}\underset{\sim}{\alpha})$. Hence, from (5.27), the detail appears as though imaged by a hypothetical telescope whose pupil function is the product of two projections of the actual telescope pupil function $Q(\underset{\sim}{\xi})$ onto a horizontal plane above the telescope at the mean turbulence altitude $\bar{h}$. The reduced effective diameter of this effective aperture causes a reduced diffraction limit and hence a broadening of the detail of $ff(\underset{\sim}{\alpha})$ in $s_{LA}(\underset{\sim}{x})$, which becomes more pronounced for a given $\bar{h} > 0$ as $|\underset{\sim}{\alpha}|$ increases. The reduction of diameter is not uniform for all directions relative to $\underset{\sim}{\alpha}$, which produces a directional dependence of spatial frequency attenuation due to partial isoplanatism [cf. Korff et al 1975, Schneiderman and Karo 1978a].

5.3.3.2 Computer Simulations of Speckle Interferometry

Figure 5.5 shows speckle interferometry spectra formed in one-dimensional computer simulations from the same object and under the same partially isoplanatic conditions as were the shift-and-add images of Figure 5.2. The autocorrelation of the object of Figure 5.2a is equal to

$$ff(\alpha) = (1+a^2)\,\delta(\alpha) + a\left(\delta(\alpha-\theta) + \delta(\alpha+\theta)\right) \tag{5.28}$$

in object space, with $\theta = -b/\ell$ and $a = 0.5$. From (5.25) and (5.28), $S_{LA}(u)$ is given by

$$S_{LA}(u) \approx (1+a^2)\,\ell^2\left(|H_{LE}(u)|^2 + \sigma(0)\,Q_4(\bar{\lambda}\ell u,0)\right) + 2a\cos(2\pi u\ell\theta)\,\ell^2\left(|H_{LE}(u)|^2 + \sigma(\theta)\,Q_4(\bar{\lambda}\ell u,\bar{h}\theta)\right). \tag{5.29}$$

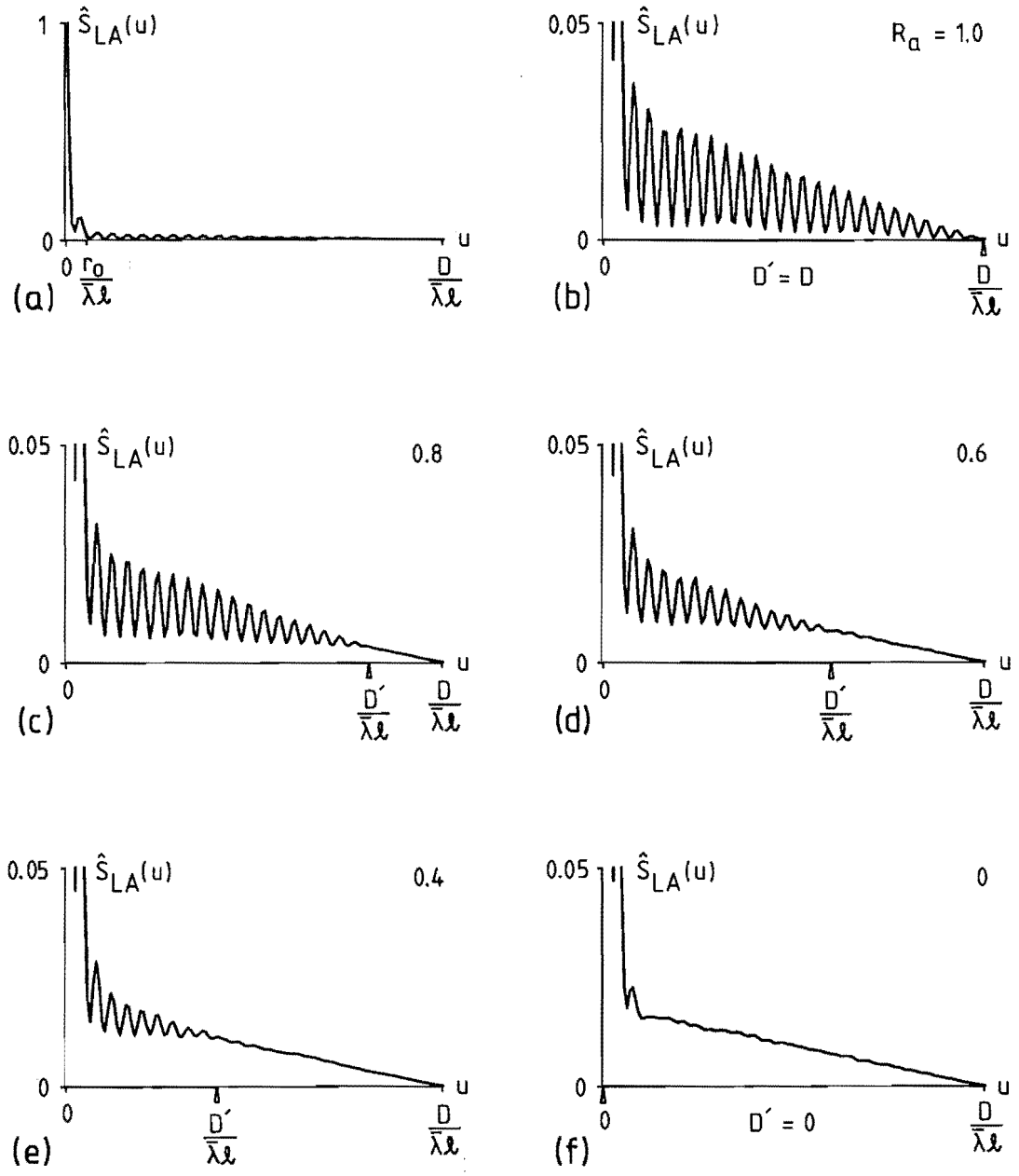


Figure 5.5: Normalised speckle interferometry spectra formed from partially isoplanatic simulated speckle images of the object shown in Figure 5.2a, under the same conditions as in Figure 5.2.

(a) Spectrum for $R_a = 1.0$ showing the relative magnitudes of low and high spatial frequency components.

(b) - (f) Spectra for $R_a = 1.0$ (b), 0.8 (c), 0.6 (d), 0.4 (e), 0 (f), plotted on an expanded vertical scale.

The speckle images were generated as described in the caption to Figure 5.2, except that here M = 1000.

In the one-dimensional simulations the turbulence is modelled as a single phase screen at height h, so that $\Delta h = 0$ and $\bar{h} = h$ (cf. ϕ2.6), and hence $\sigma(\theta) = \sigma(0) = \sigma$. The one-dimensional aperture is modelled as unaberrated and unapodised (4.5b), so that

$$\begin{aligned} Q_4(\xi,h\theta) &= \mathbf{A}\left[\mathrm{rect}\left(\frac{\xi}{D}\right)\mathrm{rect}\left(\frac{\xi+h\theta}{D}\right)\right] \\ &= \mathbf{A}\left[\mathrm{rect}\left(\frac{\xi+h\theta/2}{D-|h\theta|}\right)\right] \quad \text{for } |h\theta| < D;\ 0 \text{ otherwise} \\ &= \mathbf{A}\left[\mathrm{rect}\left(\frac{\xi}{D-|h\theta|}\right)\right] \quad \text{for } |h\theta| < D;\ 0 \text{ otherwise} \\ &= H_D'(u) \end{aligned} \tag{5.30}$$

where $H_D'(u)$ is the diffraction-limited transfer function of a telescope with a one-dimensional aperture of diameter

$$D' = \max[D - |h\theta|, 0] = D R_a \tag{5.31}$$

(cf. 4.14). For $\theta = 0$ one has

$$Q_4(\xi,0) = H_D(u) \tag{5.32}$$

with diameter D. Hence, from (5.29), Roddier et al's [1982] theory predicts that the speckle interferometry spectrum (Figure 5.5) is given by

$$\begin{aligned} S_{LA}(u) \approx\ & (1+a^2)\,\ell^2\left(|H_{LE}(u)|^2 + \sigma H_D(u)\right) \\ & + 2a\cos(2\pi u\ell\theta)\,\ell^2\left(|H_{LE}(u)|^2 + \sigma H_D'(u)\right). \end{aligned} \tag{5.33}$$

In this subsection the concern is with the extent of the second term of (5.33) in Fourier space, which is given theoretically by $D'/\bar{\lambda}\ell$ using (5.31) and in the simulations by the extent of the oscillatory segment of the $S_{LA}(u)$ visible in Figure 5.5. The predictions of (5.31) are marked with arrow-heads in Figures 5.5b-f, where evidently they match quite accurately with the extent of the oscillatory segment of the simulated $S_{LA}(u)$ (the few cycles of oscillation present in Figure 5.5f come from $|H_{LE}(u)|^2$ in the second term of (5.33)). This shows that Roddier et al's [1982] theory, and also that of Lohmann and Weigelt 1979 eq.12], describes reasonably accurately the effects of nonisoplanatism on the spatial resolution of speckle interferometry

for the one-dimensional simulations.

5.3.3.3 Comparison of Shift-and-Add with Speckle Interferometry

The magnitudes of the visibilities of the shift-and-add images of Figure 5.2 are shown in Figure 5.6. Like the $S_{LA}(u)$, each $S_{SA}(u)$ is the sum of two components, one originating from the primary peak of $f(x)$ and the other from the secondary peak of $f(x)$. The oscillatory behaviour of $S_{SA}(u)$ visible in Figure 5.6 arises from the interference between these two components. Comparing Figure 5.6 with Figure 5.5 reveals that the extent of the oscillatory segment of $S_{SA}(u)$ varies with R_a in a manner indistinguishable from that of $S_{LA}(u)$. This shows that, at least in these simulations, the loss of resolution away from the origin of the shift-and-add image due to partial isoplanatism occurs to the same degree that it does for speckle interferometry, and is therefore also well described by the theory discussed in Section 5.3.3.1. Assuming that the same is true in the real-world, the tentative conclusion is that partial isoplanatism affects shift-and-add and speckle interferometry (and hence speckle holography) in much the same way. Bagnuolo [1984] has shown theoretically that, for a single phase screen model of the atmosphere, the loss of contrast of detail away from the origin in $s_{SA}(\underset{\sim}{x})$ occurs according to the same law (5.6, 5.7; 2.38) as it does for $s_{LA}(\underset{\sim}{x})$. The simulations reported in this subsection empirically provide a corresponding result for the loss of resolution of the detail.

The degree to which the loss of resolution implied by (5.25) is apparent in practice depends on how rapidly the degree of isoplanatism or height ratio $R_h(\underset{\sim}{\alpha})$ declines compared to the decline of the diameter in Fourier space of $Q_4(\bar{\lambda}\ell\underset{\sim}{u},\bar{h}\underset{\sim}{\alpha})$, as $|\underset{\sim}{\alpha}|$ increases. For an aperture of simple shape, e.g. linear in one dimension or circular in two dimensions, consideration of (5.27) shows that the diameter of $Q_4(\bar{\lambda}\ell\underset{\sim}{u},\bar{h}\underset{\sim}{\alpha})$ is smallest in the direction parallel to $\underset{\sim}{\alpha}$, and is given in $\underset{\sim}{\xi}$-space by

$$D' = \max[D - |\bar{h}\underset{\sim}{\alpha}|, 0] \qquad (5.34a)$$

and in $\underset{\sim}{u}$-space by

$$D'/\bar{\lambda}\ell = \max[D - |\bar{h}\underset{\sim}{\alpha}|, 0] \; / \; \bar{\lambda}\ell. \qquad (5.34b)$$

The value of either of these diameters relative to their value for $|\underset{\sim}{\alpha}| = 0$ is described by the resolution ratio

$$R_r(\underset{\sim}{\alpha}) = D'/D = \max[1 - |\bar{h}\underset{\sim}{\alpha}|/D, 0]. \qquad (5.35)$$

$R_r(\underset{\sim}{\alpha})$ (5.35) and $R_h(\underset{\sim}{\alpha})$ (2.34) are compared in Figure 5.7 for a telescope with a two-dimensional circular aperture of 5m diameter and with the mean turbulence height $\bar{h}$ equal to 10km. Two curves for $R_h(\underset{\sim}{\alpha})$ are shown, one (b) for thin atmospheric turbulence with $\Delta h = 0$ (i.e. a single phase screen as in the computer simulations), and the other (c) for a more realistic case of thick turbulence with the atmospheric isoplanatic angle θ_{oA} (2.40) equal to 10 arc seconds (ϕ2.6). With the single phase screen (b), $R_h(\underset{\sim}{\alpha})$ declines at its least possible rate with $|\underset{\sim}{\alpha}|$, since $\Delta h = 0$ and hence $\sigma(\underset{\sim}{\alpha}) = \sigma(\underset{\sim}{0})$ for all $|\underset{\sim}{\alpha}|$. Nevertheless it declines considerably more rapidly than $R_r(\underset{\sim}{\alpha})$ (a). With thick turbulence the decline of $\sigma(\underset{\sim}{\alpha})$ (2.33) causes $R_h(\underset{\sim}{\alpha})$ to decline even more rapidly compared to $R_r(\underset{\sim}{\alpha})$. This means that, according to Roddier et al's [1982] theory, the loss of resolution due to partial isoplanatism is small down to a comparatively low degree of isoplanatism, with $R_r(\underset{\sim}{\alpha})$ being greater than 0.7 in the worst case (Figure 5.7d), and perhaps much closer to 1 in realistic seeing (Figure 5.7e) for $R_h(\underset{\sim}{\alpha})$ as low as 0.4. These resolution losses are very small compared to the loss of resolution in the long exposure image for a large telescope and typical seeing (ϕ2.4). Lohmann and Weigelt's [1979 ϕ3] astronomical observations provide practical confirmation that the loss of resolution due to partial isoplanatism is small (cf. the comment at the end of the penultimate paragraph of Lohmann and Weigelt [1979 ϕ2]). The resolution loss due to partial isoplanatism can thus be expected to be only a minor effect for degrees of isoplanatism as low as (say) 0.4, so that the simple mathematical model of partial isoplanatism introduced in Section 5.3.2, which predicts no resolution loss at all, can be considered to be a valid approximation to the more developed theory for $R_h(\underset{\sim}{\alpha})$ as low as 0.4.

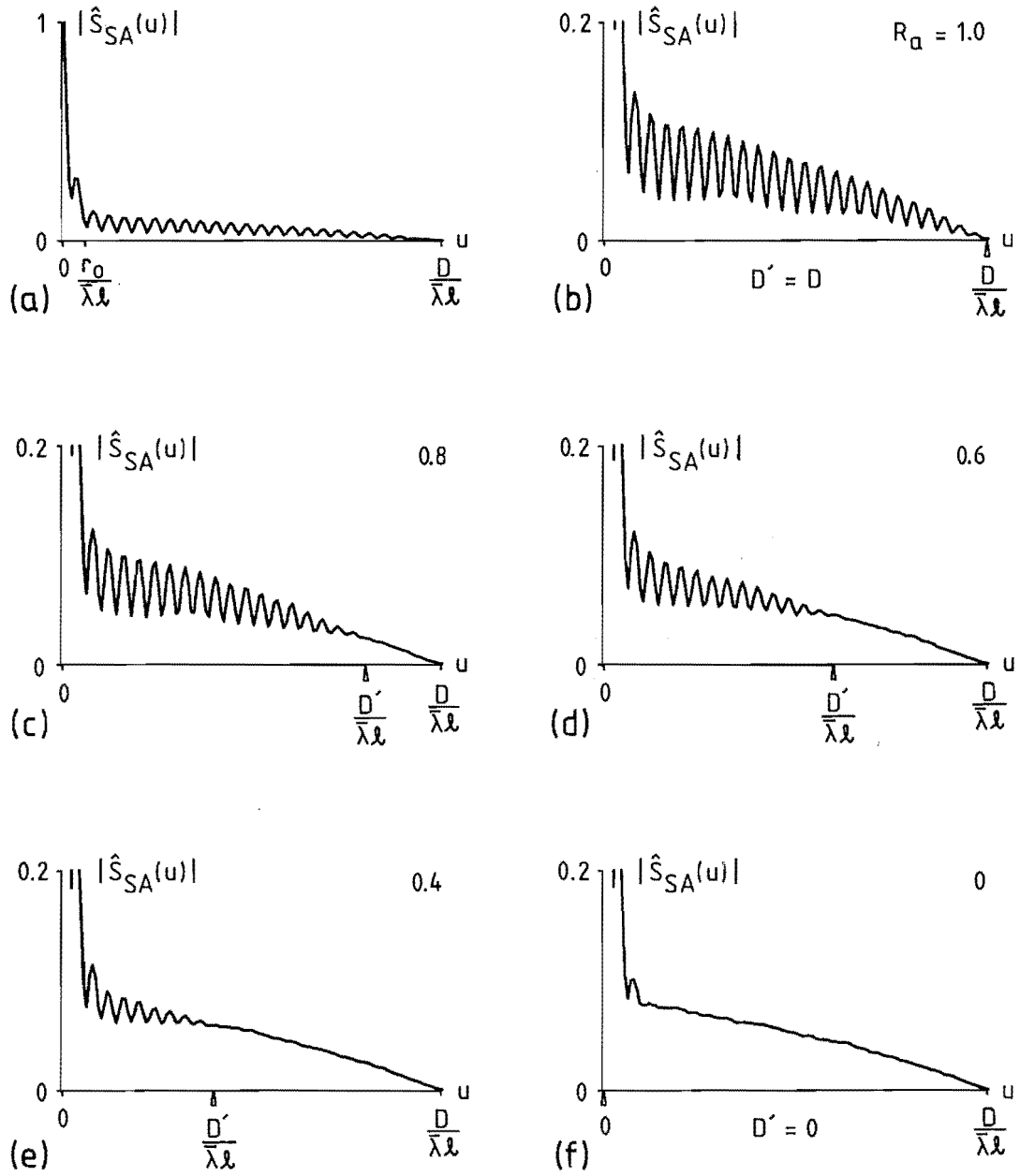


Figure 5.6: Normalised visibilities of the shift-and-add images of Figure 5.2 (the magnitudes of the visibilities are shown).

(a) Visibility of Figure 5.2b showing the relative magnitudes of low and high spatial frequency components.

(b) - (f) Visibilities of Figures 5.2b-f respectively, for which R_a = 1.0 (b), 0.8 (c), 0.6 (d), 0.4 (e), 0 (f), plotted on an expanded vertical scale.

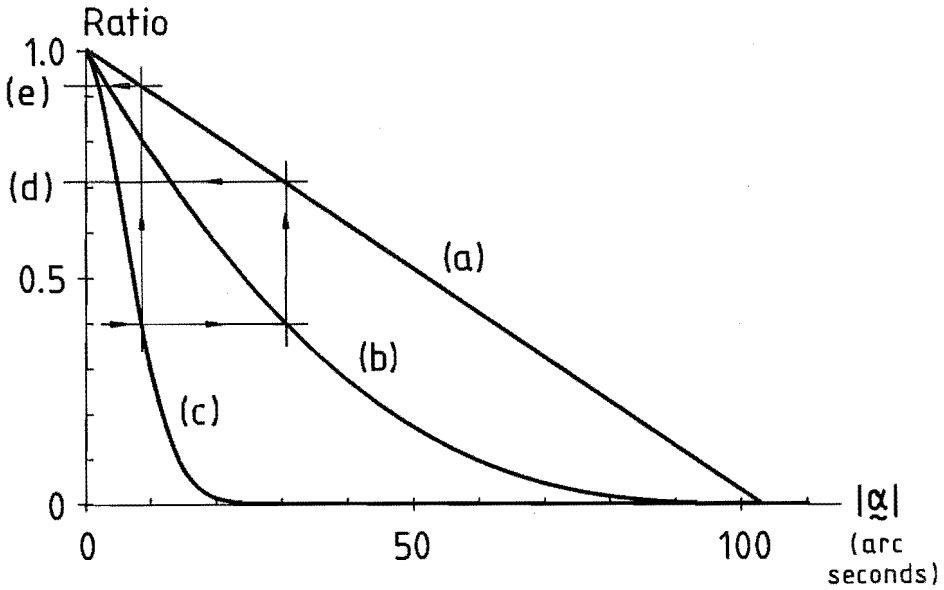


Figure 5.7: Theoretical dependence of the ratios $R_r(\underset{\sim}{\alpha})$ and $R_h(\underset{\sim}{\alpha})$ (see text) on $|\underset{\sim}{\alpha}|$, for a telescope with a two-dimensional circular aperture of 5m diameter and mean turbulence height $\bar{h}$ = 10km.

(a) Resolution ratio $R_r(\underset{\sim}{\alpha})$ (5.35).

(b) Height ratio $R_h(\underset{\sim}{\alpha})$ (2.34) for turbulence concentrated in a single layer at height $\bar{h}$.

(c) Height ratio $R_h(\underset{\sim}{\alpha})$ for thick turbulence with $\bar{h}$ = 10km and atmospheric isoplanatic angle θ_{oA} = 10 arc seconds.

(d) Resolution ratio accompanying $R_h(\underset{\sim}{\alpha})$ = 0.4 for the single layer turbulence of (b).

(e) Resolution ratio accompanying $R_h(\underset{\sim}{\alpha})$ = 0.4 for the thick turbulence of (c).

5.3.4 OBJECT POSSESSING A SPATIALLY DISTANT REFERENCE PART

5.3.4.1 Modification of Shift-and-Add

When the reference part of the object is separated from the other parts of the object by a great enough angular distance that it is resolved from them in the long exposure image, it forms a separate "speckle cloud", which is here called the reference speckle cloud, from the rest of the object in each speckle image (cf. the discussion of large-field speckle holography in Section 3.4.1). Shift-and-add can then be straightforwardly modified so that the reference part need not be brighter than the rest of the object for shift-and-add to be able to form a faithful image of the object. The modification to shift-and-add is merely to search only the region of each speckle image containing the reference speckle cloud for the position $\underset{\sim}{x}_m$ of its brightest pixel, instead of searching the entire speckle image. The speckle images are then shifted and added in the usual manner according to (5.14). The reference part of the object need only be bright enough for the reference speckle cloud to be at least a few times brighter than those parts of the speckle clouds from other parts of the object which encroach upon it. If this condition is met then $\underset{\sim}{x}_m$ depends almost entirely upon the form of the reference speckle cloud and is essentially independent of the rest of the speckle image. If the reference part of the object is also unresolved, then the central peak of the shift-and-add point spread function is of diffraction-limited extent and the shift-and-add image is essentially free of ghosts (cf. ϕ3.6.2) from either the reference part or the rest of the object.

5.3.4.2 A Computational Example

A one-dimensional computational example demonstrating the modified shift-and-add algorithm of the preceding subsection under conditions of complete and partial isoplanatism is presented in Figure 5.8. The object, Figure 5.8a, consists of an unresolved part (on the left) and a part containing resolved detail (on the right). The left hand part represents a bright unresolved star and the right hand part

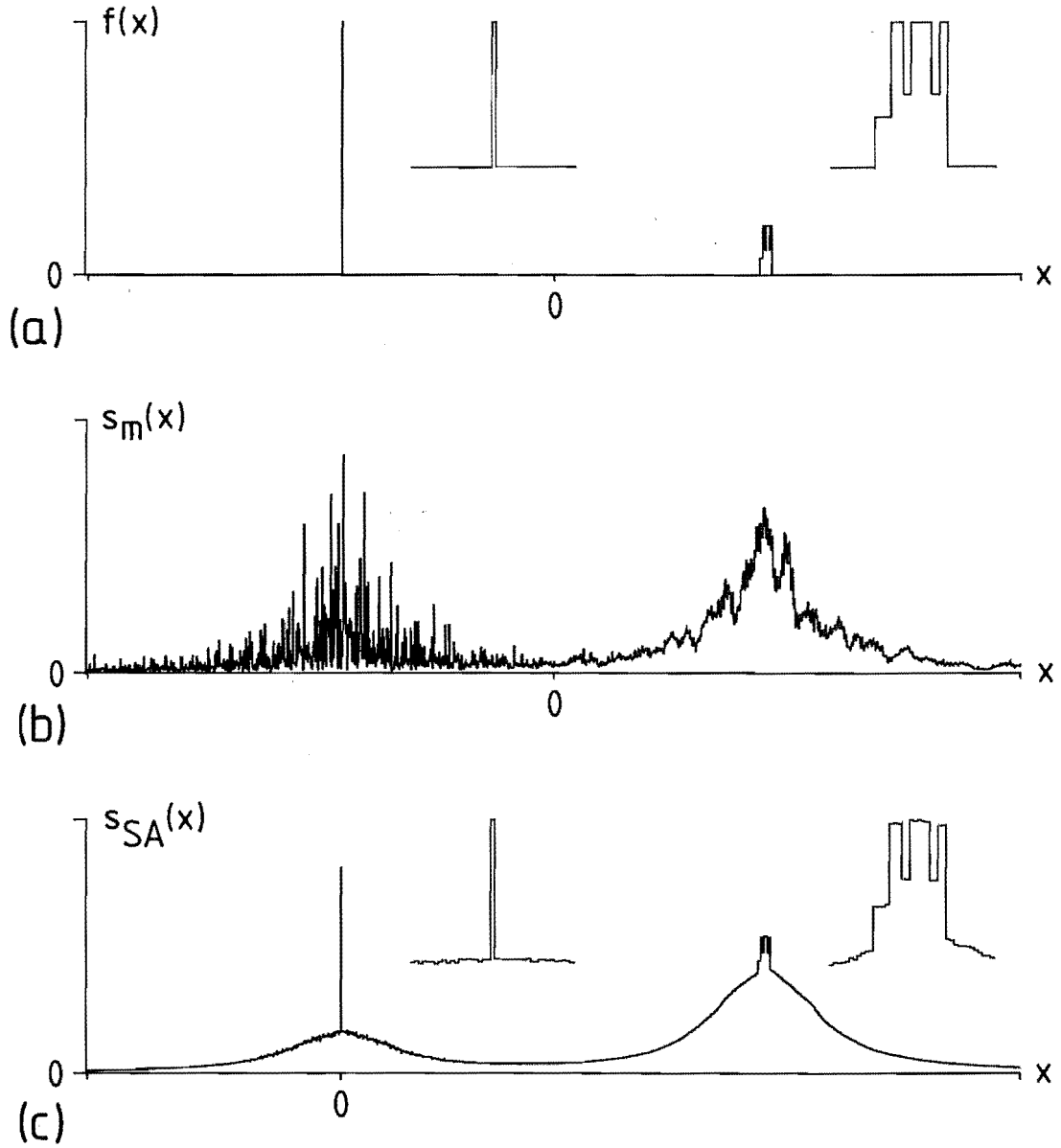


Figure 5.8: One-dimensional computer simulation of shift-and-add in completely isoplanatic and partially isoplanatic conditions with an object having a spatially distant reference part.

(a) Object.

(b) Typical isoplanatic speckle image of (a).

(c) Shift-and-add image formed from 200 speckle images similar to (b) with only the region $x < 0$ of each speckle image being searched for the position of its brightest point.

(d) Typical nonisoplanatic, contaminated speckle image of (a). The degree of isoplanatism R_h between the two parts of (a) is approximately 0.5. Uniformly distributed pseudo-random noise has been added to each pixel.

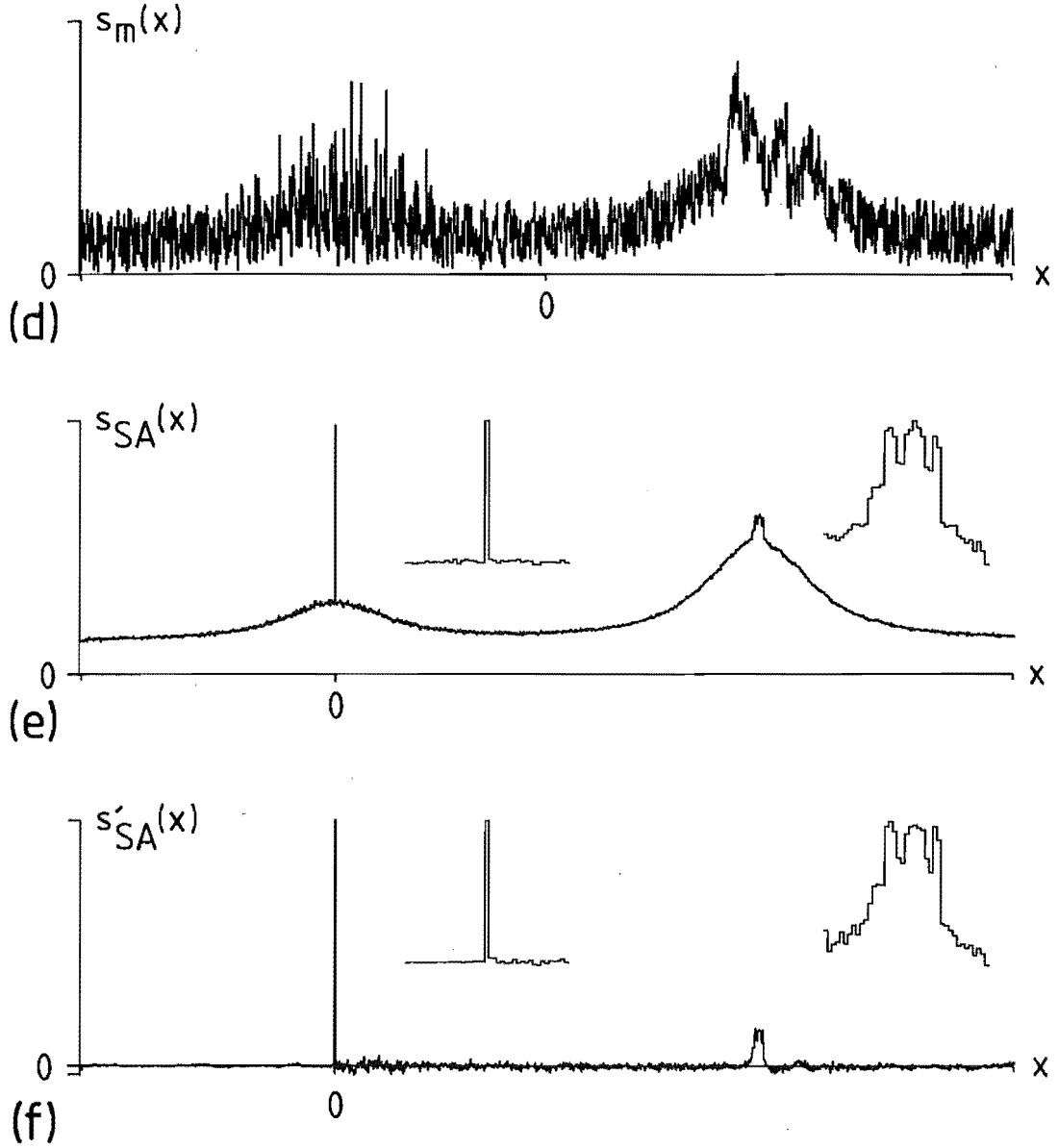


Figure 5.8 (continued):

(e) Shift-and-add image formed from 200 speckle images similar to (d) with only the region $x < 0$ of each speckle image being searched for the position of its brightest point.

(f) Defogged version of (e) formed by adding the region of (e) for $x \leq 0$ to the region of (e) for $x < 0$ reflected about $x = 0$, and deconvolving the resulting image from (e) by Wiener filtering.

Figures (b) - (e) are displayed on an arbitrary but equal vertical scale. The insets have no fixed vertical scale. The speckle images were generated with SIZE = 2048, PR = 10, CL = 100, D = SIZE, h as defined by R_a = 1.0 (b) or 0.71 (d), θ = b = 640 pixels. 1408 pixels of image space are displayed.

represents detail lying within the same isoplanatic patch as the star. Figure 5.8b shows a simulated one-dimensional speckle image of the object formed under perfectly isoplanatic conditions (cf. §4.4). The effective diameter of the seeing disc is somewhat smaller than the separation of the two parts of the object. Figure 5.8c shows the shift-and-add image formed from 200 statistically independent speckle images similar to Figure 5.8b with only the left hand speckle cloud of each speckle image being searched for the position of its brightest point, i.e. with the left hand part of the object serving as the reference part for shift-and-add. The two parts of the object are imaged with their correct relative heights on top of two separated humps of fog in Figure 5.8c (cf. insets to Figures 5.8a,c).

Because of the large angular separation between the reference part of the object and the rest of the object, the left hand hump and peak of Figure 5.8c together constitute a good estimate of the form of $h_{oSA}(x)$, the shift-and-add point spread function for an unresolved object. The fog in $s_{SA}(x)$ can thereby be suppressed by deconvolution (3.140) using only the information present in $s_{SA}(x)$. Separate observations of an unresolved reference object are not required.

The typical diameter of the isoplanatic patch (§2.6) is not very many times greater than the typical diameter of the seeing disc (§2.4). Speckle images of the type of object considered in this and the preceding subsection are therefore likely to be significantly nonisoplanatic in practice. Figures 5.8d-f illustrate the performance of the modified shift-and-add algorithm under conditions less ideal than in Figures 5.8b,c. Figure 5.8d shows a simulated nonisoplanatic speckle image generated as described in Section 4.5, with the area ratio chosen using (5.6, 5.7) to give a degree of isoplanatism R_h of 0.5 between the two parts of the object. A uniformly distributed pseudo-random number with an average value equal to the height of the fog at the origin of Figure 5.8c has also been added to each pixel of Figure 5.8d to simulate additive contamination of the speckle image by (say) sky background radiation. The shift-and-add image formed as for Figure 5.8c from 200 statistically independent versions of Figure 5.8d is shown in Figure 5.8e. Both parts of the object appear on top of the fog with no apparent loss of resolution. The detail in the right hand part of Figure 5.8e is attenuated by

approximately one half compared to Figure 5.8c, due to the partial isoplanatism of the speckle images used to form Figure 5.8e. The additive contamination to the speckle images merely raises the fog level throughout Figure 5.8e. Figure 5.8f shows the result of deconvolving from Figure 5.8e an estimate of the shift-and-add point spread function constructed from the portion of Figure 5.8e to the left of and including x = 0. The fog has been largely suppressed leaving a background of low level noise. The two parts of the object are present in Figure 5.8f with approximately their correct forms (see insets). The object detail on the right hand side of Figure 5.8f appears on a narrow hump which presumably corresponds to $s_f'(\underset{\sim}{x})$ in (5.20). This example demonstrates that shift-and-add can recover object detail close to the diffraction limit even when the speckle images are contaminated to a considerable degree by nonisoplanatism and additive noise.

5.3.5 SHIFT-AND-ADD APPLIED TO ALPHA SCORPII

The red supergiant star α^1 Scorpii (Antares) has an unresolved companion star α^2 Sco at an apparent angular distance from α^1 Sco of approximately 3 arc seconds [cf. Christou 1985a ϕ5.4]. The angular separation is such that (a) the two components of α Sco are resolved from each other in long exposure images, and (b) the degree of isoplanatism between them is likely to be significantly less than 1 but significantly greater than 0. The apparent angular diameter of α^1 Sco has been determined by Michelson stellar interferometry (ϕ2.8) to be approximately 40 milliseconds of arc [Hanbury Brown 1974 p.25]. It is thus resolvable (though only barely) by existing large optical telescopes. Its companion α^2 Sco is indubitably unresolvable. These properties make α Sco a useful object for testing the performance of the shift-and-add proposal of Section 5.3.4 on actual astronomical data.

In this subsection, the results of applying shift-and-add to approximately 2000 speckle images of α Scorpii recorded by E.K. Hege and J.M. Beckers of the Steward Observatory [Hege 1984] are reported and discussed. The speckle images were recorded using the Steward Observatory speckle camera [Hege et al 1982b] with one mirror of the

Multiple Mirror Telescope [cf. Hege et al 1985]. The received light was filtered to a centre wavelength of 470nm and a bandwidth of 10nm. A neutral density filter with a transmission of 13% attenuated the light from α^1 Sco to make the brightness of its image more nearly equal to that of α^2 Sco.

5.3.5.1 Data Processing

The speckle images captured with the Steward Observatory speckle camera were sent to the University of Canterbury on digital magnetic tape as frames of 256 × 128 pixels with 8 bits per pixel. There were 1960 useable speckle images on the tapes. Geometrical distortion from the speckle camera was rectified from each raw speckle image [cf. Castleman 1979 ϕ8] using the coordinates of the image of a rectangular grid of points supplied by Hege [1984]. Speckle processing was then applied to the 1960 rectified speckle images.

In each speckle image, the images of α^1 and α^2 Scorpii appeared as separate speckle clouds, with the speckle cloud from α^2 Sco located within the left hand 128 × 128 pixels and the speckle cloud from α^1 Sco located within the right hand 128 × 128 pixels. A long exposure image was formed by adding together all of the speckle images. Shift-and-add images were formed in two different ways. The left-shift-and-add image was formed by searching only the left hand 128 × 128 pixels of each speckle image for the position $\underset{\sim}{x}_{lm}$ of its brightest pixel, and computing

$$s_{lSA}(\underset{\sim}{x}) = \langle s_m(\underset{\sim}{x} + \underset{\sim}{x}_{lm})\rangle_M. \tag{5.36}$$

The right-shift-and-add image was formed by searching only the right hand 128 × 128 pixels of each speckle image for the position $\underset{\sim}{x}_{rm}$ of its brightest pixel, and computing

$$s_{rSA}(\underset{\sim}{x}) = \langle s_m(\underset{\sim}{x} + \underset{\sim}{x}_{rm})\rangle_M. \tag{5.37}$$

When more than one pixel shared the maximum brightness in the region searched, the first such pixel encountered during the search was taken

to be the brightest pixel.

The left-shift-and-add image corresponds to the computational example of Section 5.3.4.2. Its shift-and-add reference (α^2 Sco) is unresolved, and the right hand part of $s_{lSA}(\underset{\sim}{x})$ is therefore an image of the red supergiant α^1 Sco unaffected by ghosting but affected by partial isoplanatism. In the right-shift-and-add image, the image of α^2 Sco instead of α^1 Sco is affected by partial isoplanatism. Comparing $s_{rSA}(\underset{\sim}{x})$ with $s_{lSA}(\underset{\sim}{x})$ allows the effects of partial isoplanatism on each of the shift-and-add images to be assessed.

5.3.5.2 Results and Discussion

A single speckle image of α^1 and α^2 Scorpii (after rectification of geometrical distortion) is shown in Figure 5.9a. This image was selected from a sample of about twenty speckle images to demonstrate partial isoplanatism between the two components of α Sco. Partial isoplanatism is apparent to the eye in the overall bottom left to top right diagonal arrangement of speckles in each part of the speckle image and in the relative positioning of clumps of three or four bright speckles in each part. In most of the speckle images of the sample, some similarity in the arrangement of groups of speckles in each part was noticeable, though less so than in Figure 5.9a. The long exposure image formed from the sum without shifts of all of the speckle images is shown in Figure 5.9b. The two parts of α Sco are well resolved from each other. Note that α^1 Sco appears considerably brighter than α^2 Sco despite its image having been attenuated by a neutral density filter. This difference in brightness affects the relative brightnesses of the two parts in each of the shift-and-add images. Figure 5.9c shows the left-shift-and-add image formed using α^2 Sco as the shift-and-add reference, and Figure 5.9d shows the right-shift-and-add image formed with α^1 Sco as the shift-and-add reference. The ratio of brightness between the two parts of each shift-and-add image is quite different in Figures 5.9c and d. This is clearly a result of partial isoplanatism. If α^1 Sco is treated as unresolved, the modelling of Section 5.3.2.2 together with a comparison of the relative brightnesses of each part of Figures 5.9b-d yields the value $R_h = 0.5 \pm 0.1$ for the degree of isoplanatism

in this observation across the angular separation (≈ 3 arc seconds) of the two components of α Sco.

The relative angular diameters of the images of α^1 and α^2 Scorpii in Figures 5.9c and d are difficult to assess from the photographs because of the different brightnesses of the images. To facilitate comparisons of angular diameter, radially averaged profiles of each part of Figures 5.9b-d are plotted on a normalised vertical scale in Figure 5.10. The brightest pixel of each part of Figures 5.9b-d was chosen as the centre point for its radial average. In Figure 5.10 the dotted lines are the profiles for α^2 Sco and the solid lines are the profiles for α^1 Sco.

The profiles from the long exposure image are virtually identical in form (Figure 5.10a). For the left-shift-and-add image, the profile from α^2 Sco (Figure 5.10b, dotted line) is considerably narrower than the profile from α^1 Sco (solid line). The greater width of the image of α^1 Sco is due partly to α^1 Sco actually having a larger apparent angular diameter than α^2 Sco, and partly to the partial isoplanatism existing between α^1 and α^2 Sco (cf. ∮5.3.3). In $s_{rSA}(\underset{\sim}{x})$, partial isoplanatism should alter the relative widths of α^1 Sco and α^2 Sco oppositely to its effect in $s_{lSA}(\underset{\sim}{x})$, whereas the greater angular

Figure 5.9 (facing page): Astronomical speckle observations of the stars α^1 and α^2 Scorpii. α^2 Sco is on the left and α^1 Sco (the red supergiant Antares) is on the right. Their angular separation is approximately 3 arc seconds.

(a) A selected speckle image showing visually apparent partial isoplanatism between the speckle clouds from α^1 and α^2 Sco.

(b) Long exposure image.

(c) Left shift-and-add image (α^2 Sco, arrowed, served as the shift-and-add reference).

(d) Right-shift-and-add image (α^1 Sco, arrowed, served as the shift-and-add reference).

1960 speckle images similar to (a) were used to form each of the images (b), (c), and (d).

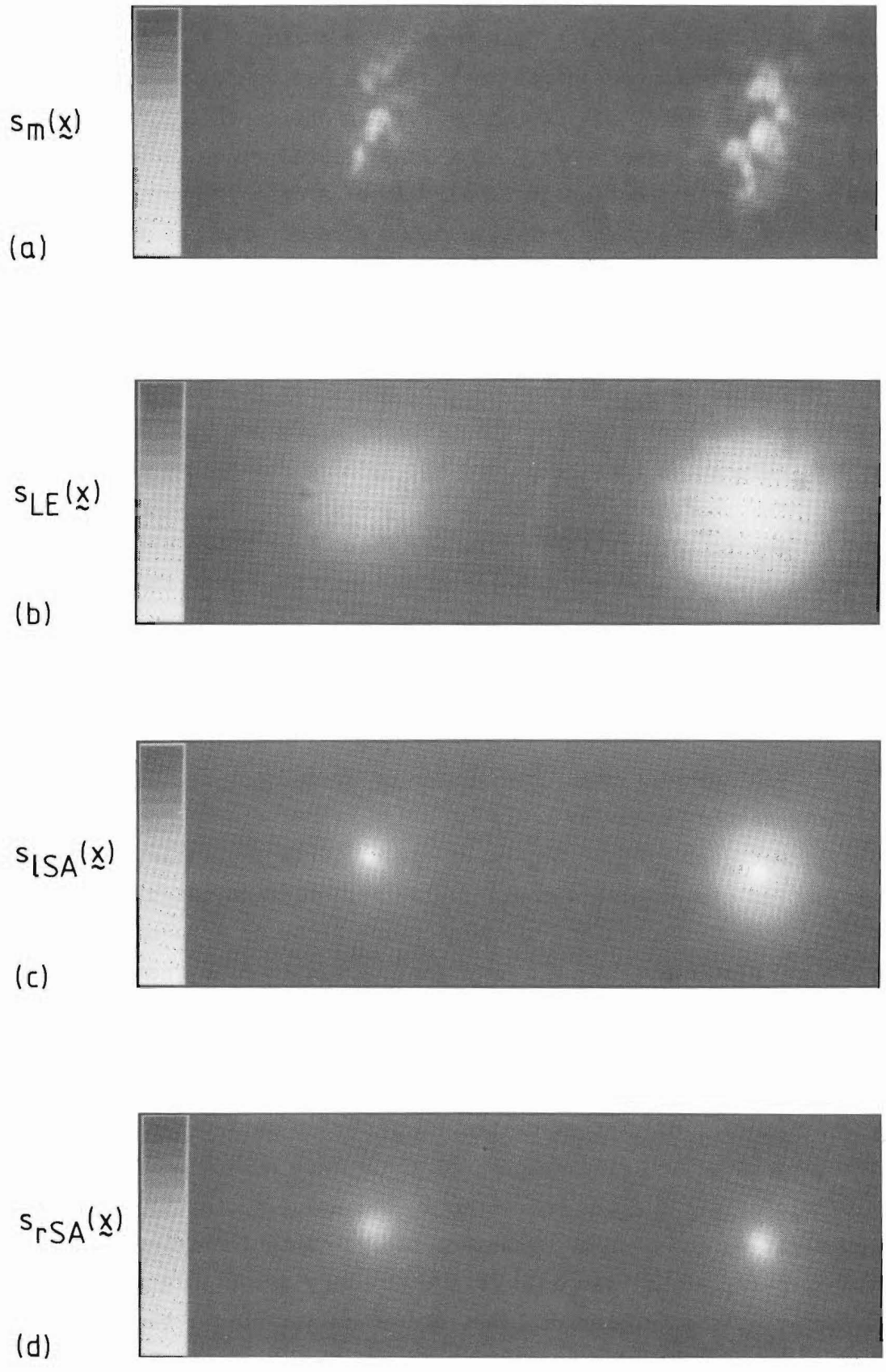

$s_m(\underset{\sim}{x})$
(a)
$s_{LE}(\underset{\sim}{x})$
(b)
$s_{lSA}(\underset{\sim}{x})$
(c)
$s_{rSA}(\underset{\sim}{x})$
(d)

Figure 5.9

diameter of α^1 Sco should tend to make its image relatively wider in both $s_{rSA}(\underset{\sim}{x})$ and $s_{lSA}(\underset{\sim}{x})$. This provides a means of distinguishing the broadening effects of partial isoplanatism and the different apparent angular diameters of the two stars. The profiles of $s_{rSA}(\underset{\sim}{x})$ are shown in Figure 5.10c. The profile of α^2 Sco (dotted line) is broader than the profile of α^1 Sco (solid line), but by a relative amount that is noticeably less than the relative difference of width of the profiles in Figure 5.10b. This must be because the apparent angular diameter of α^1 Sco is greater than that of α^2 Sco. To make this comparison more obvious, the profiles from the parts of α Sco used as shift-and-add references in $s_{lSA}(\underset{\sim}{x})$ and $s_{rSA}(\underset{\sim}{x})$ are shown together in Figure 5.10d, and the profiles from the parts of α Sco not used as shift-and-add references in $s_{lSA}(\underset{\sim}{x})$ and $s_{rSA}(\underset{\sim}{x})$ are shown together in Figure 5.10e. The profile from α^1 Sco is slightly wider than the profile from α^2 Sco in both of these figures, which is a clear indication of α^1 Sco's greater apparent angular diameter. Note that the parts of $s_{lSA}(\underset{\sim}{x})$ and $s_{rSA}(\underset{\sim}{x})$ whose profiles appear in Figure 5.10d correspond to shift-and-add images of α^2 Sco alone and of α^1 Sco alone, formed independently in the conventional manner (ϕ3.6.2).

The apparent angular diameter of α^1 Scorpii is too small relative to the resolving power of a single MMT mirror for any significant detail of α^1 Sco to be deduced from the images of Figure 5.9, or for its angular diameter to be reliably estimated from Figure 5.10. Attempts to enhance the right hand parts of Figures 5.9c and d by Wiener filtering the left hand part of Figure 5.9c from them were unsuccessful. The diffraction-limited point spread function of a single MMT mirror, computed from (3.5) with $\bar{\lambda}$ = 470nm and $P(\underset{\sim}{\xi})$ defined by the geometry of the MMT mirror's aperture (whose diameter is approximately 1.8m) [Strittmatter 1977], has a half-intensity diameter of approximately 55 milliseconds of arc. This conveniently defines a resolution limit for the MMT mirror. The angular diameter of α^1 Scorpii as measured by Michelson stellar interferometry is approximately 40 milliseconds of arc [Hanbury Brown 1974 p.25], which is less than the resolution limit of the MMT mirror. Thus, both components of α Sco are unresolved by one mirror of the MMT according to a conventional interpretation of resolution.

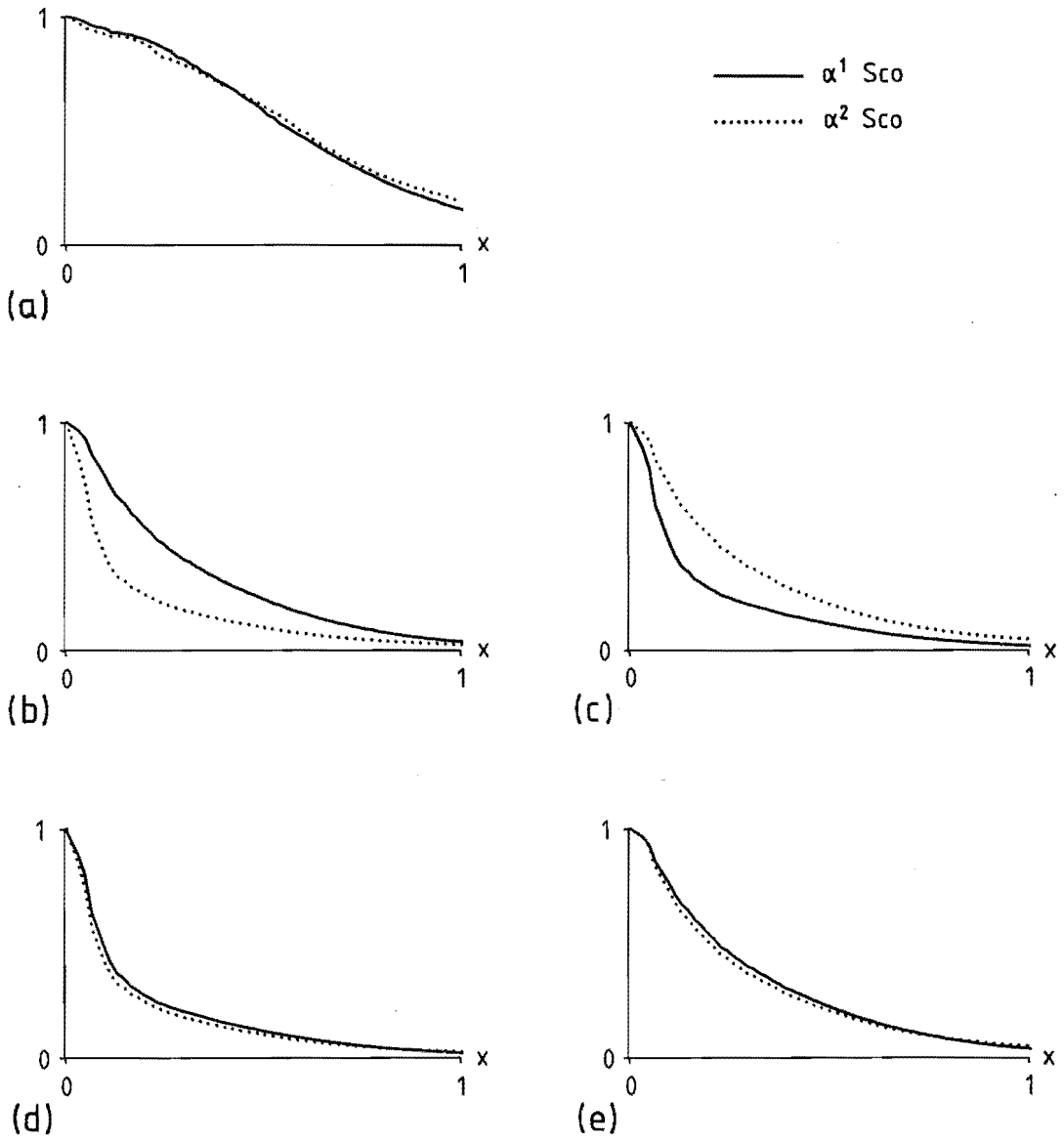


Figure 5.10: Radially averaged profiles of brightness in the images of α^1 and α^2 Scorpii shown in Figure 5.9. The solid line is α^1 Sco and the dotted line is α^2 Sco in each figure. Each profile is normalised to unit brightness at x = 0. The horizontal scale is in arbitrary units.

(a) Long exposure image (Figure 5.9b left and right).

(b) Left-shift-and-add image (Figure 5.9c left and right).

(c) Right-shift-and-add image (Figure 5.9d left and right).

(d) Shift-and-add reference parts (Figure 5.9c left, Figure 5.9d right).

(e) Shift-and-add non-reference parts (Figure 5.9c right, Figure 5.9d left).

Two anomalous aspects of the results presented in Figure 5.10 deserve comment. First, there are no sidelobes (Airy rings) visible in the profiles of either α^1 or α^2 Scorpii in Figure 5.10d, whereas previous experience with shift-and-add indicates that sidelobes should be present in the shift-and-add image of an unresolved object where the central peak of the image meets the fog (cf. Figure 3.10a, Christou et al [1986 Figure 1]). Christou [1985b] has remarked on a similar lack of sidelobes in certain weighted shift-and-add images of γ Orionis formed with narrowband optical filters. There may be some systematic aberration which has affected the data set processed here as well as the ones noted by Christou [1985b]. Secondly, the amount of broadening of the shift-and-add profiles due to nonisoplanatism (cf. Figures 5.10b,c) appears to be greater than the computer simulations of Sections 5.3.3 and 5.3.4 and the theory of Roddier et al [1982] (ϕ5.3.3, Figure 5.7) predict for the observed value of R_h. This could be due to faulty rectification of geometrical distortion from the speckle images, or to some systematic degradation of the data set as noted above, or to Roddier et al's [1982] theory not applying to shift-and-add images in practice. The available information is insufficient to decide among these alternatives.

5.3.6 CONCLUSIONS

The computer simulations of shift-and-add with partially isoplanatic speckle images described in Sections 5.3.1 and 5.3.4, and the simple mathematical model of partial isoplanatism presented in Section 5.3.2, show that partial isoplanatism causes an attenuation of detail in the shift-and-add image and a slight loss of resolution of the detail, with no other untoward effects. The simulations confirm Bagnuolo's [1984] theoretical expression for the attenuation of the detail with a single phase screen atmosphere. Roddier et al's [1982] theory for speckle interferometry accurately predicts the loss of resolution in the computer simulations. The simulations suggest that the effects of nonisoplanatism on shift-and-add and speckle interferometry are very similar.

The loss of resolution due to partial isoplanatism in shift-and-add images formed from actual astronomical data (ϕ5.3.5) is

greater than expected from Roddier et al's [1982] theory for the observed degree of isoplanatism, though it is still much less than the resolution loss in the long exposure image due to the seeing. The discrepancy may arise from inapplicability of Roddier et al's [1982] theory to shift-and-add in real conditions, invalidity of Roddier et al's modelling of the speckle images for this observation, or faulty data processing.

Shift-and-add processing of astronomical speckle images of two stars, α^1 and α^2 Scorpii, separated by about 3 arc seconds on the celestial sphere, gives an estimate of 0.5 ± 0.1 for the degree of isoplanatism R_h between the two stars. This figure is consistent with measurements of R_h for other double stars with a 1.5m telescope by Ebersberger and Weigelt [1985]. A reliable image or an estimate of the apparent angular diameter of the red supergiant α^1 Sco could not be formed because the angular diameter of α^1 Sco is too small relative to the resolution of the 1.8m aperture used to gather the speckle images. To obtain a useful image or an estimate of the angular diameter of α^1 Sco by shift-and-add, it will be necessary to process data collected with a larger aperture, for example the 6.9m aperture of the six mirrors of the MMT operating together coherently [Hege et al 1985]. Nevertheless, the results presented in Section 5.3.5 show that shift-and-add is capable of revealing small differences in the angular diameters of celestial objects even if the objects are unresolved in the conventional sense (e.g. smaller than the Rayleigh limit). The profiles presented in Figure 5.10e show that the small difference in angular diameter discernible in Figure 5.10d is still evident to approximately the same degree in the presence of broadening due to partial isoplanatism. This suggests that accurate characterisation of the broadening would allow its effects to be effectively removed from the shift-and-add image.

Nonisoplanatism appears to degrade the shift-and-add images of objects (which would be formed faithfully under perfectly isoplanatic conditions) in a benign manner, in contrast to the reported effects of nonisoplanatism on the Knox-Thompson method (ϕ3.5.2) [Karo and Schneiderman 1975, Fried 1979]. This, combined with the computational simplicity of shift-and-add, appears to make it a viable technique for forming high spatial resolution images of regions of the celestial

sphere lying within the same isoplanatic patches as unresolved stars bright enough to form good shift-and-add reference sources.

5.4 SHIFT-AND-ADD WITH PHOTON-LIMITED SPECKLE IMAGES

Speckle images are formed with short exposure times and (usually) narrow optical bandwidths (ϕ2.5), compared to conventional long exposure astrophotography. For this reason, speckle images of the majority of interesting astronomical objects are significantly photon-limited (cf. ϕ3.3.3). The behaviour of shift-and-add under photon-limited conditions is therefore an important subject of study. This section presents the results of one-dimensional computer simulations of shift-and-add with photon-limited speckle images.

5.4.1 CHANGES IN THE SHIFT-AND-ADD IMAGE

Figure 5.11 shows a series of shift-and-add images (3.135) formed from computer simulated photon-limited speckle images (ϕ4.6) of the object depicted in Figure 3.9a. The average number of photons per speckle image, $\bar{N}_p$, ranges from 10,000 in Figure 5.11a to 3 in Figure 5.11h. A non-photon-limited speckle image (effectively with $\bar{N}_p = \infty$) of the same object with the same simulated seeing conditions appears in Figure 3.9h. To take account of the possibility of non-unique brightest pixels in the photon-limited speckle images, the shift-and-add processing strategy was modified from the simple shift-and-add prescription of choosing the brightest pixel to shift to the centre of image space, as is explained in Section 5.4.2 below. The shift strategy 1 defined in Section 5.4.2 was invoked when generating the results presented in Figure 5.11.

An immediately obvious effect of photon limiting is the generation of a narrow spike, which following Christou et al [1985] is here called the photon spike, on top of the primary peak of the shift-and-add images shown in Figure 5.11. The spike is almost negligible for $\bar{N}_p \geq 10{,}000$ (compare Figure 5.11a with Figure 3.9h), but grows relative to the rest of the shift-and-add image as $\bar{N}_p$ decreases. In the simulated speckle images, each photon impact was modelled as being

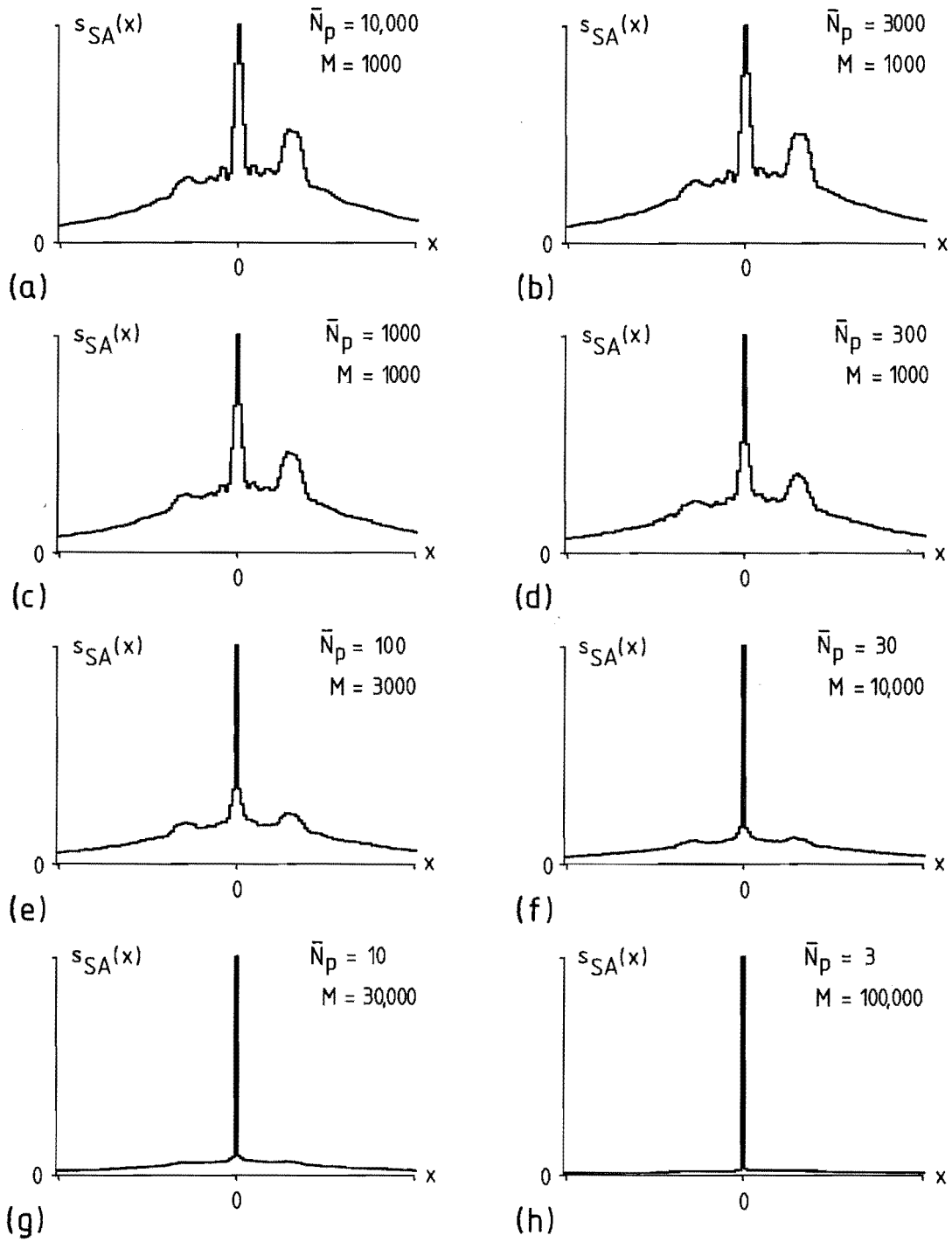


Figure 5.11: Shift-and-add with one-dimensional simulated photon-limited speckle images.

(a) - (h) Shift-and-add images formed from M photon-limited speckle images of Figure 3.9a with $\bar{N}_p$ photons on average per speckle image.

The speckle images were generated with SIZE = 256, PR = 15, CL = 100, D = SIZE/4, h = 0, $\bar{N}_p$ and M as indicated in the figures. SIZE/2 pixels are displayed.

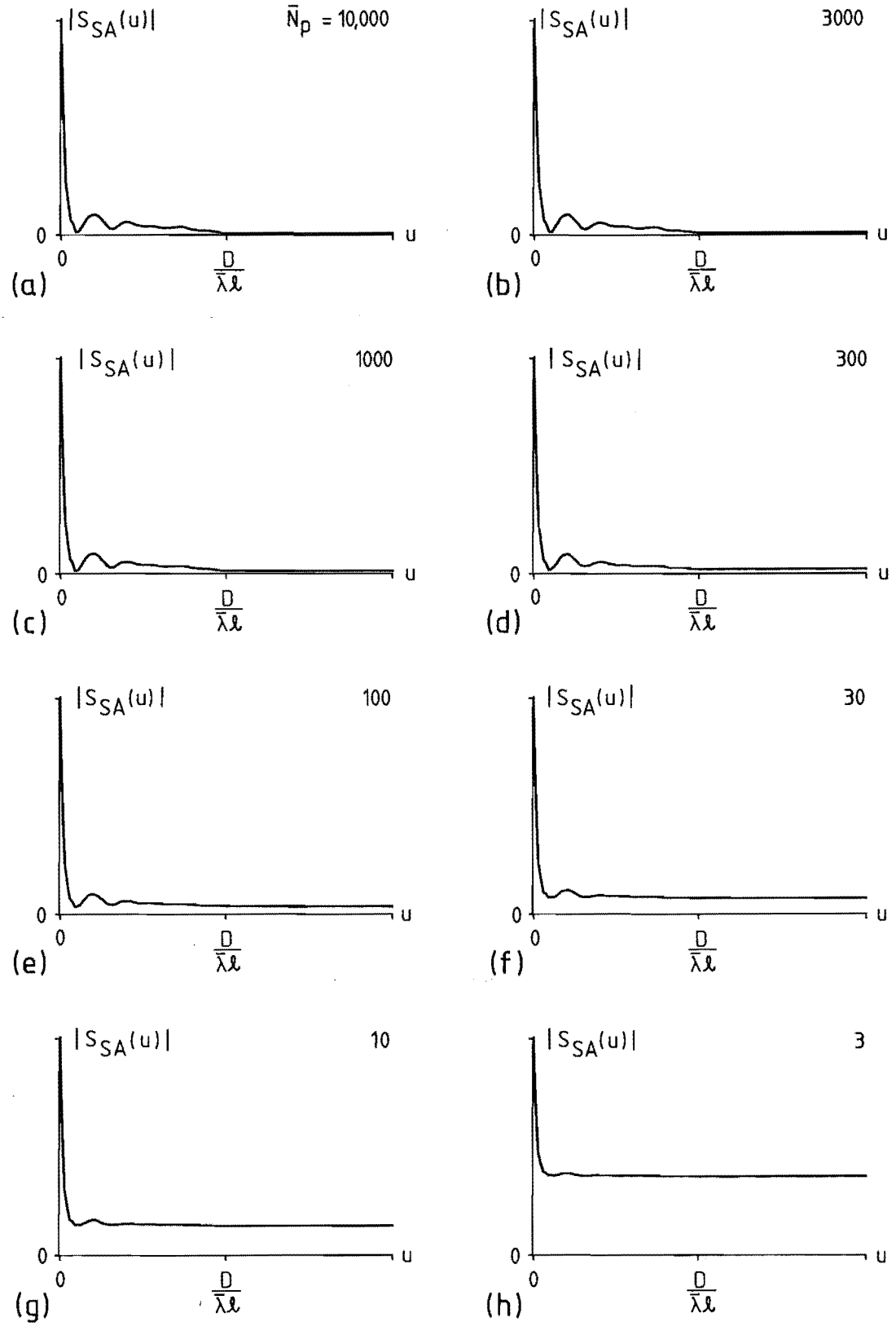


Figure 5.12: Visibilities of the shift-and-add images of Figures 5.11a-h (the visibility magnitudes are displayed). The photon bias due to photon limiting gives the visibilities nonzero value above the diffraction limit.

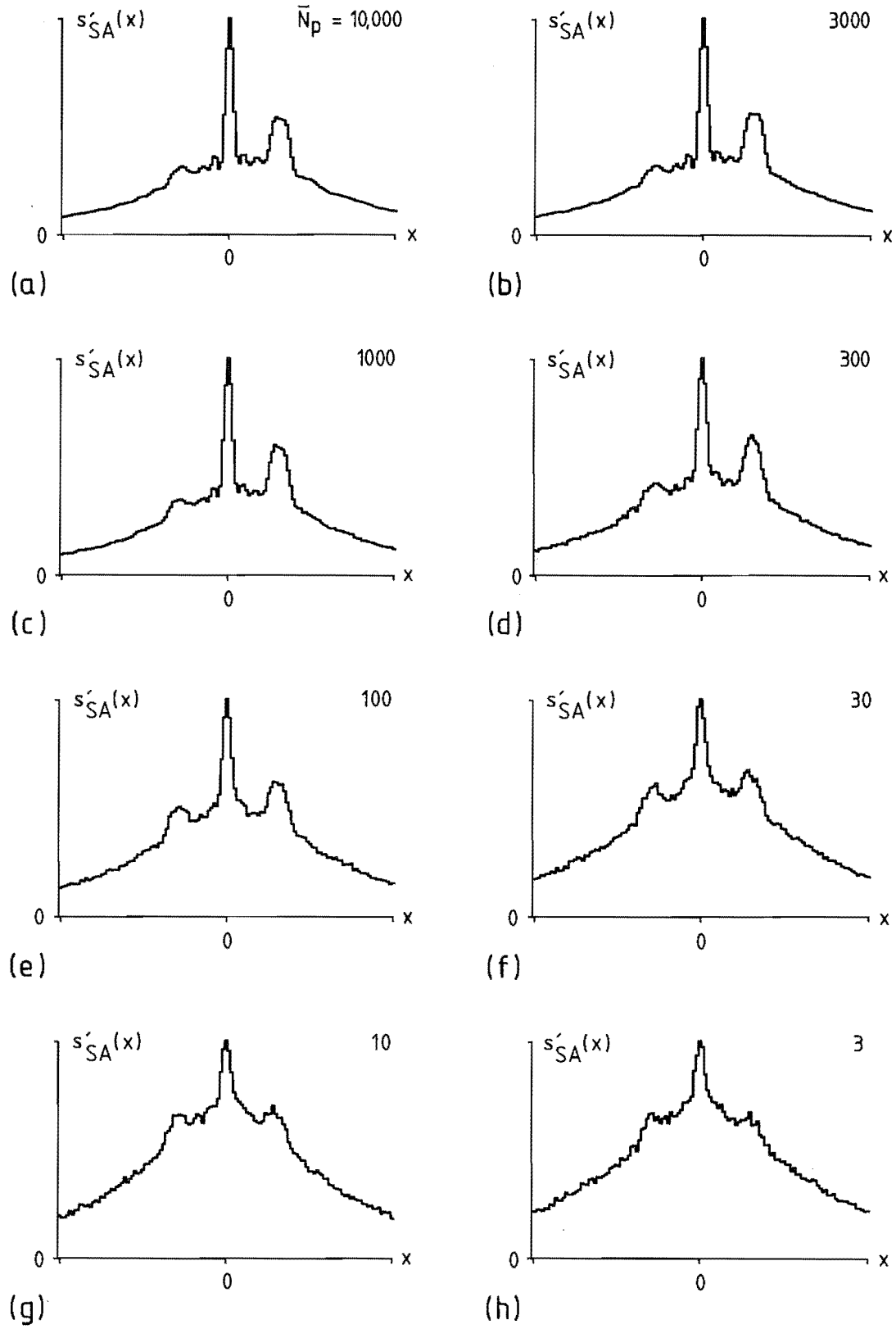


Figure 5.13: Despiked versions of the shift-and-add images of Figures 5.11a-h. The despiking was performed in the Fourier domain by subtracting a constant equal to the photon bias from the visibilities (Figure 5.12) of the shift-and-add images.

detected at a single pixel of image space, so that the photon spike in the images of Figure 5.11 is only one pixel wide.

In the Fourier domain (see Figure 5.12), the photon spike appears as a photon bias added to the spectral content of the rest of the shift-and-add image, similar to the photon bias which appears with speckle interferometry under photon-limited conditions (ϕ3.3.3). As with speckle interferometry, the photon bias in shift-and-add extends beyond the diffraction limit of the telescope if the detector's spatial frequency response is wider than that of the telescope. In Figure 5.12 the photon bias is uniform since (in the simulations whose results are presented in this section) the detector point spread function was modelled as occupying a single pixel. Christou et al [1985, 1986] describe the photon bias in weighted shift-and-add images (ϕ3.6.3) formed from astronomical data and demonstrate that it can be straightforwardly removed in practice if the detector psf can be modelled accurately. The removal is effected by fitting a model of the detector psf to the shift-and-add visibility beyond the diffraction limit and then subtracting the fitted psf from the shift-and-add visibility. Christou et al call this procedure despiking.

In the examples presented in Figures 5.11 and 5.12, the photon bias is uniform and so despiking merely involves subtracting a real-valued constant from the spectra to set them to zero above the diffraction limit. The despiked shift-and-add images so obtained are shown in Figure 5.13. For $\bar{N}_p \geq 1000$ the despiked images are virtually identical to the shift-and-add image formed with the same seeing under non-photon-limited conditions (Figure 3.9h). For $\bar{N}_p < 1000$, however, changes in the form of the despiked images become apparent. In comparison to the height of the primary peak above the fog (cf. Figure 5.1c), the secondary peak reduces in height, the ghost peak increases in height, and the fog level increases. According to the theory reviewed in Section 3.3.3, the form of the averaged autocorrelation of speckle images (3.14) produced by speckle interferometry is unaffected by decreasing $\bar{N}_p$ apart from the growth of a photon spike and a decrease in signal-to-noise ratio for any fixed number M of speckle images. Figure 5.13 demonstrates that, unlike speckle interferometry, the shift-and-add image experiences changes of form additional to the growth of a photon spike and a decrease in signal-to-noise ratio as

$\bar{N}_p$ decreases.

5.4.2 THE INFLUENCE OF SHIFT STRATEGY

Simple shift-and-add (ϕ3.6.2) requires the brightest pixel of each speckle image to be identified so that the speckle image can be translated to shift this pixel to the centre of image space. This pixel is called the shift pixel in what follows. Since each pixel of a photon-limited speckle image is quantised to a finite number of levels, there is a nonzero probability that more than one pixel is "brightest". This probability increases as the average number of photons detected per speckle image decreases. When a speckle image has non-unique brightest pixels, the shift strategy by which shift-and-add selects the shift pixel must be elaborated from the simple prescription of choosing the brightest pixel of the speckle image. The properties of the chosen shift strategy have an increasing influence on the form of the shift-and-add image as $\bar{N}_p$ is reduced, because the proportion of speckle images having non-unique brightest pixels (requiring invocation of the elaborate shift strategy) increases as $\bar{N}_p$ decreases.

An optimum shift strategy for photon-limited conditions should minimise the sensitivity of the form of the shift-and-add image to changes in $\bar{N}_p$ when $\bar{N}_p$ is small, and should also minimise ghosting when $\bar{N}_p$ is large. Here, five particular shift strategies are considered:

1. MEDIAN SHIFTING: select the median brightest pixel as the shift pixel, i.e. the brightest pixel which, for N brightest pixels, is the $(N-1)/2^{th}$ brightest pixel from the leftmost (or rightmost) one. If N is even then randomly select the $N/2^{th}$ or the $(N-2)/2^{th}$ brightest pixel from the leftmost (or rightmost) one with equal probabilities.

2. NEAREST-CENTROID SHIFTING: select the brightest pixel nearest to the centroid of all the brightest pixels as the shift pixel. The centroid x_c of N brightest pixels at positions x_n is defined as

$$x_c = \frac{1}{N} \sum_{n=1}^{N} x_n \, . \qquad (5.38)$$

If two brightest pixels are equally nearest to the centroid, randomly select either one with equal probabilities.

3. LEFTMOST SHIFTING: select the leftmost of the brightest pixels as the shift pixel.

4. RIGHTMOST SHIFTING: select the rightmost of the brightest pixels as the shift pixel.

5. RANDOM SHIFTING: randomly select the shift pixel from the position of any of the detected photons with equal probability for each photon, i.e. randomly choose any pixel in the speckle image with probability proportional to the value of the speckle image at that pixel.

Shift strategies 1 to 4 reduce to the simple shift-and-add strategy described in Section 3.6.2 when the brightest pixel is unique. They represent initial approximations to the ideal of minimising ghosting when $\bar{N}_p$ is large. Shift strategies 3 and 4 cause an arbitrary directional bias in the shift-and-add image and hence may not be generally applicable in practice, but they are included as geometrically extreme ways of choosing among the brightest pixels. Shift-and-add with shift strategy 5 implements a probabilistic formulation of the averaged autocorrelation of the speckle images and is therefore equivalent to speckle interferometry, yielding an estimate of the autocorrelation of the object instead of the object itself. Shift strategies 2 and 5 are immediately generalisable to two or more dimensions.

Figures 5.14 and 5.15 summarise the performance of shift-and-add with the five shift strategies described above in one-dimensional computer simulations. One-dimensional simulated photon-limited speckle images of a binary object with unresolved parts (cf. §5.2) were generated as described in Section 4.6. The object was defined by (5.1b) with a = 0.5 and b = 20 pixels (x_0 is arbitrary). Four statistically independent ensembles of photon-limited speckle images were generated for each of several different values of $\bar{N}_p$. Shift-and-add was performed on each ensemble with each of the five shift strategies listed above. For Figure 5.14 the shift-and-add images were despiked as described in Section 5.4.1. The despiked shift-and-

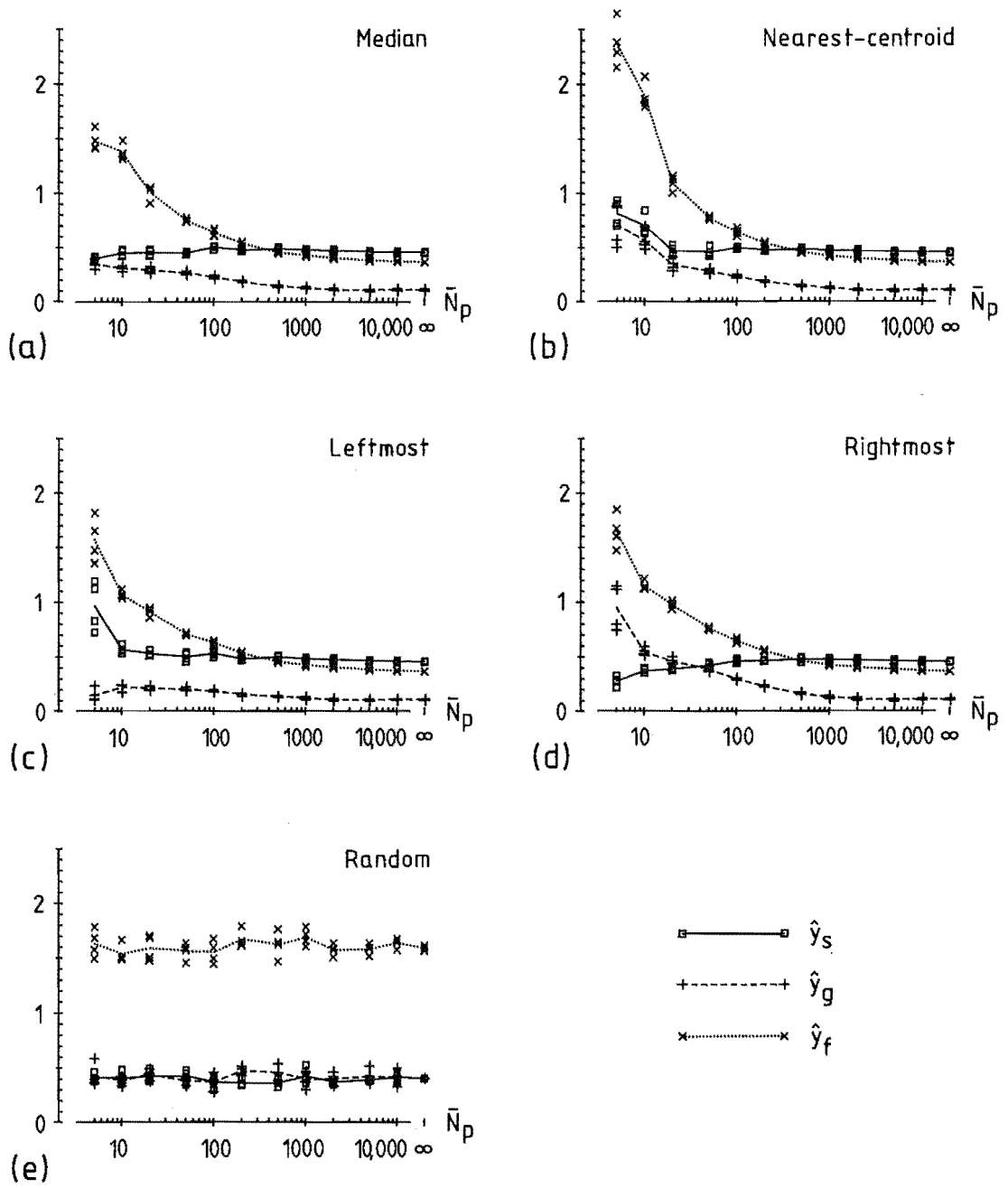


Figure 5.14: Effects of photon limiting on the form of the despiked shift-and-add image, in one-dimensional simulations with the five shift strategies described in the text. Each graph shows the variation with $\bar{N}_p$ of the normalised heights of the secondary peak ($\hat{y}_s$), the ghost peak ($\hat{y}_g$), and the fog ($\hat{y}_f$) (cf. Figure 5.1) of the despiked images. The object was binary with unresolved parts, with a = 0.5 and b = 20. The speckle images were generated with SIZE = 256, PR = 15, CL = 100, D = SIZE/4, h = 0, $\bar{N}_p$ as indicated, M ranging from 1000 for $\bar{N}_p$ = 10,000 to 40,000 for $\bar{N}_p$ = 5.

add images are characterised in Figure 5.14 by the normalised heights of their primary peak, secondary peak, and fog (ϕ5.2). The graphs in Figures 5.14a-e show the variation of these quantities with $\bar{N}_p$ for each of the five shift strategies. Figure 5.15 shows the variation of the photon bias (i.e. the value of $|S_{SA}(u)|$ above the diffraction limit) with $\bar{N}_p$ for the five shift strategies.

The results summarised by Figures 5.14 and 5.15 have the following notable features:

(a) $\hat{y}_s$, $\hat{y}_g$, and $\hat{y}_f$ (Figure 5.14) remain essentially constant with $\bar{N}_p$ only for random shifting. For the other shift strategies they vary with $\bar{N}_p$, as noted for median shifting at the end of Section 5.4.1.

(b) The photon bias (Figure 5.15) is essentially constant with $\bar{N}_p$ only for random shifting, for which it is approximately equal to unity. For the other shift strategies it is approximately equal to unity at the lowest $\bar{N}_p$ shown ($\bar{N}_p = 5$), and increases slowly with increasing $\bar{N}_p$. Note that the photon bias decreases relative to the rest of the shift-and-add visibility as $\bar{N}_p$ increases with all of the shift strategies, because the component of the shift-and-add visibility other than the photon bias increases with $\bar{N}_p$ approximately proportionally to $\bar{N}_p$.

(c) The trend for $\hat{y}_f$ to increase as $\bar{N}_p$ decreases is less pronounced for median shifting than for nearest-centroid shifting. Also, $\hat{y}_s$ and

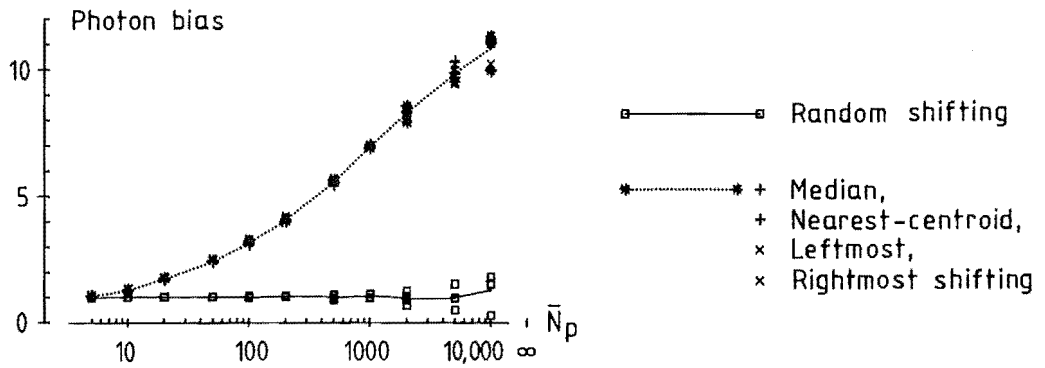


Figure 5.15: Variation of photon bias with $\bar{N}_p$ in the visibilities of the shift-and-add images whose despiked versions are the subject of Figure 5.14.

$\hat{y}_g$ both increase sharply for $\bar{N}_p < 200$ for nearest-centroid shifting but not for median shifting.

(d) With leftmost shifting, $\hat{y}_s$ rises gradually and $\hat{y}_g$ remains relatively low as $\bar{N}_p$ decreases. With rightmost shifting, $\hat{y}_s$ falls and $\hat{y}_g$ rises past $\hat{y}_s$ as $\bar{N}_p$ decreases. Note that the behaviour of the leftmost and rightmost shifting strategies would be transposed for an object equal to the mirror image of the one used here.

(e) For the severest photon limiting simulated, i.e. $\bar{N}_p = 5$, $\hat{y}_s$ and $\hat{y}_g$ are almost indistinguishable for all of the shift strategies. This result includes leftmost and rightmost shifting, since $\hat{y}_s$ for leftmost shifting approximately equals $\hat{y}_g$ for rightmost shifting, and vice versa, when $\bar{N}_p = 5$.

5.4.3 CONCLUSIONS

The most prominent effect of photon limiting on shift-and-add is the creation of a narrow photon spike at the centre of the shift-and-add image, corresponding to a photon bias in the visibility domain analogous to that encountered with speckle interferometry under photon-limited conditions (ϕ3.3.3). Christou et al [1985, 1986] have shown how the photon spike may be removed from the shift-and-add image in practice. The consequences of improperly removing the photon spike in shift-and-add should affect only the primary peak of the shift-and-add image, whereas (for example) in the Knox-Thompson method they seriously degrade the entire image [Nisenson and Papaliolios 1983]. Compensating shift-and-add for photon bias should be no more difficult than similarly compensating speckle interferometry [cf. Hege et al 1982b].

The one-dimensional computer simulations reported in this section show that the shift-and-add image experiences significant changes other than the formation of a photon spike when the photon limiting is severe, e.g. less than 1000 photons per speckle image in the simulations. The nature and severity of the changes depend on the shift-and-add shift strategy used to choose among non-unique brightest pixels in photon-limited speckle images. Four of the five shift

strategies investigated here inevitably choose the brightest pixel as the shift pixel when the brightest pixel is unique. Ghosting and the fog both increase in the despiked shift-and-add image formed with these shift strategies as $\bar{N}_p$ decreases. The increase in ghosting is the more important effect because it erodes the advantage of shift-and-add over speckle interferometry of preserving some part of the visibility phase of the object. The fifth shift strategy, random shifting, preserves no visibility phase for any $\bar{N}_p$ because it produces an image equivalent to that formed by interferometry (3.14).

When $\bar{N}_p$ became small enough in the simulations ($\bar{N}_p = 5$) that nearly all of the speckle images had at most one photon per speckle image, the ghost peak and the secondary peak were either nearly or completely indistinguishable in the shift-and-add images formed by each shift strategy. This suggests that shift-and-add with the shift strategies investigated here may only be able to estimate the autocorrelation of the object when there is at most one photon per pixel in the speckle image. Shift-and-add might then be best applied by increasing the optical bandwidth, exposure time, and pixel size of the speckle images until at least two photons coincide in one pixel of each speckle image. This amounts to a trade-off between spatial resolution and retention of the object's visibility phase.

Provided a sizeable proportion of speckle images have pixels registering two or more photons, shift-and-add is able to maintain a significant difference between the ghost and secondary peaks (cf. Figure 5.14a for $\bar{N}_p = 10$) even though the speckle images are severely photon-limited. Shift-and-add can then contribute useful information about the visibility phase of the object to composite speckle imaging schemes incorporating speckle interferometry and Fourier phase retrieval (cf. final paragraph of ϕ3.4.2), despite severe photon limiting.

CHAPTER 6

ZERO-AND-ADD

6.1 INTRODUCTION

This chapter describes a new speckle imaging principle called zero-and-add. Zero-and-add uses the complex zeros of the visibilities of one-dimensional speckle images to identify the fixed visibility component corresponding to the object. It is related to the use of complex zeros in consistent deconvolution [Bates et al 1976a, McKinnon et al 1976], astronomical imaging [Bates and Napier 1972, Walker 1981b, Bates and Fright 1982; cf. Bates 1982b §7.2, Dainty 1984 §7.4.4], and the Fourier phase problem in a general context (cf. §3.4.2) [Bates 1978, Bates and Fright 1984, Fright 1984 §§3.3, 3.4, 4.1]. Three papers have so far been prepared describing zero-and-add. These are Bates et al [1985], Davey et al [198_], and Sinton et al [198_].

The pertinent part of the theory of complex visibility zeros is summarised in Section 6.2. The principle of zero-and-add is presented and illustrated with an "ideal" computational example in Section 6.3. Section 6.4 describes a practical algorithm for performing zero-and-add with speckle images that have been "contaminated", as all speckle images inevitably are in practice. Section 6.5 presents computational investigations of the effects of three important types of contamination on the fidelity of the images formed by zero-and-add. Conclusions and discussion follow in Section 6.6.

6.2 COMPLEX VISIBILITY ZEROS

The complex zeros of the visibility of a function are introduced and described briefly in this section to support the description and discussions of zero-and-add presented in the following sections of this chapter. More complete treatments of complex visibility zeros are given by Bates [1969] [cf. Bates 1978], Napier

[1971 ϕ5], and Requicha [1980].

Let s(x) denote a one-dimensional image, which is a function of the one-dimensional real variable x. The visibility, or spatial Fourier transform, of s(x) is defined by (cf. 1.12a)

$$S(u) = \int_{-\infty}^{\infty} s(x)\, e^{-j2\pi ux}\, .dx, \tag{6.1}$$

where u is also a one-dimensional real variable. S(u) and s(x) are in general complex functions. It is useful to replace the real variable u in (6.1) by a complex variable

$$w = u + jv, \tag{6.2}$$

so that the visibility S(u) becomes a function

$$S(w) = \int_{-\infty}^{\infty} s(x)\, e^{-j2\pi wx}\, .dx \tag{6.3}$$

of the complex variable w. The visibility is said to be <u>analytically continued</u> into the complex w-plane. S(u) is equal to the value of S(w) along the real axis, or u-axis, of the complex plane. The relationship between S(u) and S(w) is illustrated in Figure 6.1. It is further discussed and nicely illustrated by Napier [1971 ϕ5]. In the remainder of this chapter, the term "visibility" is generalised to encompass both S(w) and S(u).

The <u>support</u> T of an image is here taken to be the smallest continuous interval $T = [x_1; x_2]$ of image space outside which the image is zero, i.e.

$$\begin{aligned} &s(x) \neq 0 \quad \text{for } x = x_1 \text{ or } x = x_2, \\ &s(x) = 0 \quad \text{for all } x < x_1 \text{ or } x > x_2. \end{aligned} \tag{6.4}$$

The <u>extent</u> of an image is the width of its support,

$$E = x_2 - x_1. \tag{6.5}$$

The visibility S(w) of an image s(x) whose extent is finite

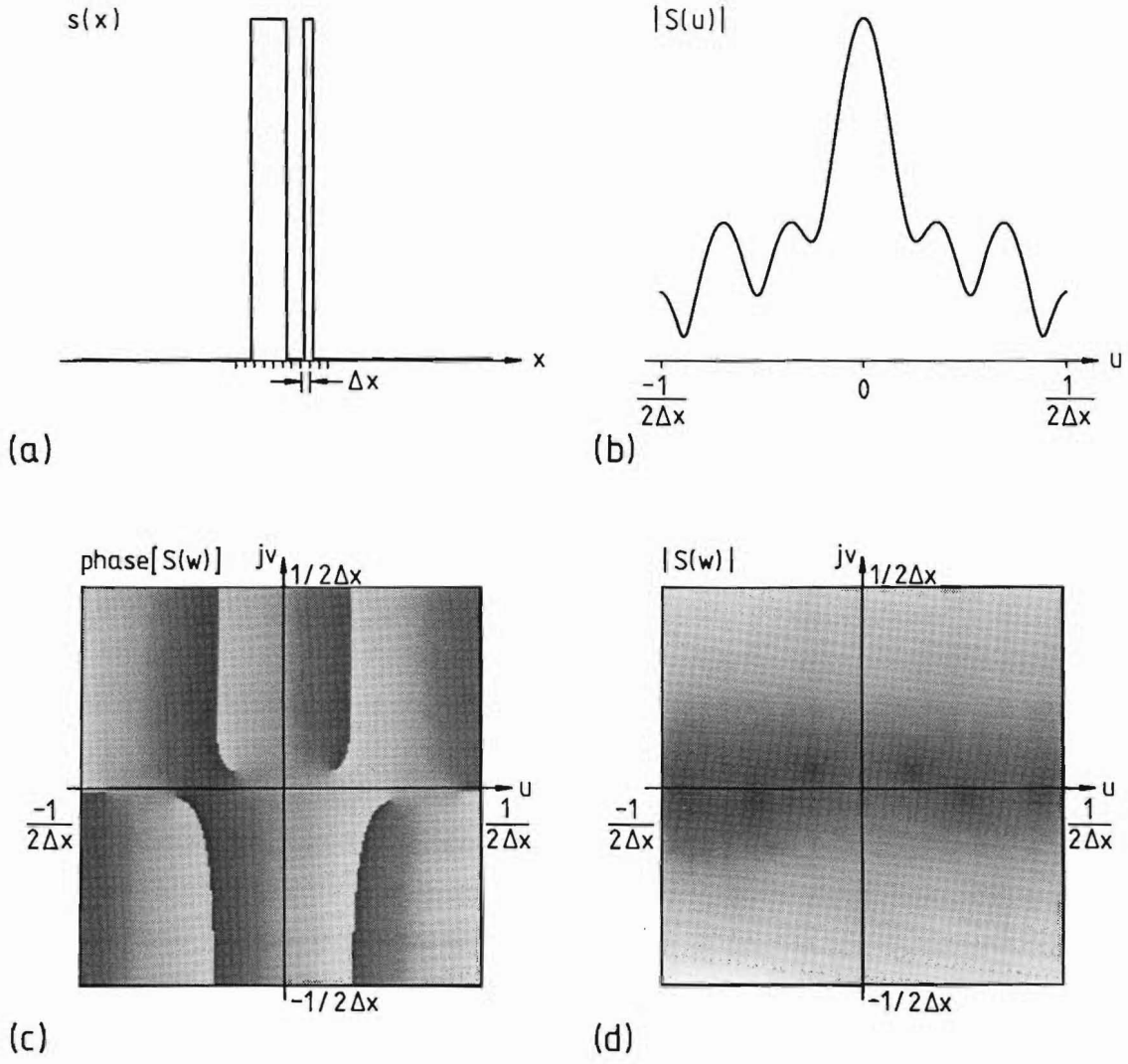


Figure 6.1: The visibility S(u) of an image s(x), and the visibility S(w) obtained by analytically continuing S(u) into the complex plane w = u + jv.

(a) A one-dimensional image s(x). Δx is the sample spacing or pixel size in image space.

(b) Magnitude of the visibility S(u) = $\mathbf{F}[s(x)]$.

(c) Phase of the analytically-continued visibility S(w).

(d) Magnitude of S(w), displayed with a logarithmic scale of intensity. (b) is a section through (d) along the u-axis.

is an entire function of exponential type [Levin 1964]. It is therefore characterised by the particular values $w = \omega_i$ of w for which S(w) is zero, i.e.

$$S(\omega_i) = 0. \tag{6.6}$$

The ω_i are here called the zeros of S(w) or the visibility zeros of s(x). Given the set $\{\omega_i\}$ of all of the ω_i, it is possible to reconstruct $C \cdot s(x)$ where C is an arbitrary complex constant.

Napier [1971 ϕ5] presents a collection of theorems on the zeros of entire functions of exponential type. Two important items in this collection stem from the following relations, verifiable immediately from (6.3):

	IMAGE		VISIBILITY		
Real	$s(x) = s^*(x)$	$\leftrightarrow$	$S(w) = S^*(-w^*)$		(6.7)
Even	$s(x) = s(-x)$	$\leftrightarrow$	$S(w) = S(-w)$	Even	(6.8)

From (6.7), the visibility zeros of a real image occurs in pairs $(\omega_i, -\omega_i^*)$ which are reflections of each other in the imaginary axis of the w-plane. From (6.8), the visibility zeros of an even image occur in pairs $(\omega_i, -\omega_i)$ which are point-symmetric about the origin of the w-plane. An image which is both real and even, such as the autocorrelation of another real image, has visibility zeros occurring in groups of four, $(\omega_i, \omega_i^*, -\omega_i, -\omega_i^*)$, which are reflections of each other in both the real and imaginary axes of the w-plane.

Another important theorem from Napier [1971 ϕ5] applies to images, such as the intensity of a field, which are non-negative. The theorem states that a non-negative image,

$$s(x) \geq 0 \quad \forall\ x, \tag{6.9}$$

can have visibility zeros on the imaginary axis of the w-plane only if it is zero for all x. This can be shown by replacing w with jv in (6.3) and noting that the exponential function in the integral then becomes purely real and positive.

The number of zeros in the set $\{\omega_i\}$ is denumerably infinite [Bates 1969]. However, only a finite subset $\{\omega_p\}$ of $\{\omega_i\}$ is required to represent s(x) to within a given accuracy. Following Bates [1969] and Fright [1984 ϕ3.3], the members of the set $\{\omega_p\}$ are here referred to as the complex zeros of S(w) or the complex visibility zeros of s(x). Some or all of the ω_p may be real, but in general they are all complex. The set $\{\omega_p\}$ is the goal of computational procedures for determining the visibility zeros of a given s(x). Bates [1969] and Bates and Napier [1972] represent s(x) by a truncated Fourier series of 2N+1 terms and obtain 2N complex zeros carrying information about s(x). Bruck and Sodin [1979] represent s(x) by a set of N+1 samples in the image domain, from which N complex zeros may be computed as described in the next paragraph. In each of these approaches the number of independent complex zeros obtained is one less than the number of independent parameters initially used to describe the image, which leads to the arbitrary scaling constant mentioned after equation (6.6) when the image is reconstructed from its complex zeros.

The approach of Bruck and Sodin [1979] for computing $\{\omega_p\}$ is convenient when the image is stored in sampled form in a digital computer. Let s(x) be the sampled function

$$s(x) = \sum_{n=0}^{N} s_n \delta(x - x_n), \tag{6.10a}$$

where s_n are the sample values and

$$x_n = n\Delta x \tag{6.10b}$$

are the sample positions. Δx is the sample spacing in the image domain. Since the image is of finite extent, the number of samples, N+1, is finite. The extent E of the sampled image is given by

$$E = (N+1)\Delta x. \tag{6.11}$$

The visibility of s(x) is obtained by substituting (6.10) into (6.3), giving

$$S(w) = \sum_{n=0}^{N} s_n \, e^{-j2\pi n \Delta x w}. \tag{6.12}$$

On defining a new complex variable

$$z = e^{-j2\pi\Delta x\, w}, \tag{6.13}$$

(6.12) becomes

$$S(w) = \sum_{n=0}^{N} s_n z^n = P(z), \tag{6.14}$$

which is the discrete z-transform [Stanley 1975 ϕ4] of the image $\{s_n\}$. $P(z)$ is a polynomial of degree N in z, whose N + 1 coefficients are equal to the image sample values. According to the fundamental theorem of algebra, $P(z)$ may be factorised into N linear factors to give

$$S(w) = P(z) = C \prod_{p=1}^{N} (z - \zeta_p) \tag{6.15}$$

[cf. Requicha 1980 eq.(2)], where C is a constant whose value is necessarily real if the image is real. The constants ζ_p in the N factors of $P(z)$ in (6.15) are identical to the N complex zeros of $P(z)$ in the complex z-plane. They can be computed from the N + 1 coefficients s_n of $P(z)$ in (6.14) using standard polynomial-factorising or zero-finding algorithms. The w-plane zeros ω_p are then found by inverting (6.13). Note from (6.13) that

$$\zeta_p = e^{-j2\pi\Delta x\, \omega_p}, \tag{6.16}$$

thereby implying that

$$\omega_p = u_p + jv_p = \frac{\mathrm{Ln}(\zeta_p)}{-j2\pi\Delta x}$$

$$= \frac{-\mathrm{phase}[\zeta_p]}{2\pi\Delta x} + j\frac{\ln|\zeta_p|}{2\pi\Delta x}, \tag{6.17}$$

where

$$\mathrm{Ln}(\zeta_p) = \ln|\zeta_p| + j\,\mathrm{phase}[\zeta_p] \tag{6.18}$$

is the complex logarithm of ζ_p and $\mathrm{phase}[\zeta_p]$ denotes the argument or phase of the complex quantity ζ_p. The image $\{s_n\}$ may be recovered

from its complex zeros $\{\omega_p\}$ by reversing the procedure described above, i.e. by computing $\{\zeta_p\}$ from $\{\omega_p\}$ using (6.16), determining the polynomial P(z) from $\{\zeta_p\}$ using (6.15), and taking the coefficients of P(z) as the image $\{s_n\}$ (cf. 6.14). All of the examples presented in the following subsections of this chapter have been computed by the procedure described in this paragraph, using the FORTRAN subroutines CPOLY [Jenkins and Traub 1972, Withers 1974] or RPOLY [Jenkins 1975] to factorise the polynomial P(z).

The relationship between the w-domain and the z-domain deserves further explanation. The mapping from w to z (6.13) is many-to-one, and the mapping from z to w (cf. 6.17) is one-to-many. The real axis of the w-plane maps to the unit circle on the z-plane. The upper half of the w-plane maps to the interior of the unit circle on the z-plane. The visibility S(w) of the sampled image s(x) consists of periodically repeated vertical strips of width $1/\Delta x$ in the w-plane, each strip corresponding to a 2π range of phase$[\zeta_p]$ in (6.17). The periodic visibility S(w) has a denumerably infinite number of zeros, comprising finite sets of N zeros per period. The N zeros whose real parts are in the range $\frac{-1}{2\Delta x} \leq u < \frac{1}{2\Delta x}$ (cf. Figure 6.1), corresponding to the principal range $[-\pi;\pi)$ of phase$[\zeta_p]$ in (6.17), are taken to form the set of complex zeros $\{\omega_p\}$. The z-domain and w-domain representations of S(w) (6.14, 6.12) are the algebraic polynomial and the trigonometric polynomial, respectively, of Requicha [1980].

6.3 THE ZERO-AND-ADD PRINCIPLE

An isoplanatic, one-dimensional speckle image $s_m(x)$ of an object f(x) is represented by the convolution

$$s_m(x) = f(x) \odot h_m(x) \tag{6.19}$$

(3.3e). The analytically-continued visibility of the speckle image is therefore given by

$$S_m(w) = F(w)H_m(w). \tag{6.20}$$

Since the speckle visibility $S_m(w)$ is the product of F(w) and $H_m(w)$,

$S_m(w)$ is zero wherever either or both of $F(w)$ or $H_m(w)$ are zero. Denoting the complex visibility zeros of $s_m(x)$, $f(x)$, and $h_m(x)$ by $\{\omega\}_{sm}$, $\{\omega\}_f$, and $\{\omega\}_{hm}$ respectively, (6.20) implies that

$$\{\omega\}_{sm} = \{\omega\}_f \cup \{\omega\}_{hm} , \tag{6.21}$$

i.e. the set of complex visibility zeros of $s_m(x)$ is the union of the sets of complex visibility zeros of $f(x)$ and $h_m(x)$. Note that the members of the three sets in (6.21) are complex numbers and that the cardinal numbers of all of the three sets are finite in practice.

The key idea behind zero-and-add is that $F(w)$ is unchanging and has the same complex zeros for all speckle images, whereas $H_m(w)$ is a random function of m and its complex zeros can be expected to have random values from one H_m to the next. The principle of zero-and-add is therefore to compare the $\{\omega\}_{sm}$ from several or many speckle images, find the particular complex zeros which are common to all of the $\{\omega\}_{sm}$ (this is the set $\{\omega\}_f$), and from $\{\omega\}_f$ compute $f(x)$ in the manner described in Section 6.2.

Since the complex visibility zeros ω of a one-dimensional image $s(x)$ are complex numbers, the set $\{\omega\}$ can be represented as a collection of points on the complex plane, whose positions represent the values of the individual members of $\{\omega\}$. When represented in this way, $\{\omega\}$ is called the zero map of $s(x)$ or of $S(w)$, and is denoted by $ZS(w)$. Zero maps are conveniently represented in a digital computer by two-dimensional arrays of pixels. An integer value associated with each pixel equals the number of complex zeros in $\{\omega\}$ lying within the rectangular region of the complex plane corresponding to that pixel. Zero maps can be subjected to the same arithmetic operations as are commonly applied to sampled representations of two-dimensional images. In particular, zero maps may be added together, subtracted from one another, and convolved with blurring functions. The relation (6.21) describing the complex visibility zeros of an isoplanatic speckle image is expressed in terms of zero maps as

$$ZS_m(w) = ZF(w) + ZH_m(w). \tag{6.22}$$

The zero-and-add principle is illustrated with the aid of zero maps in Figure 6.2. A one-dimensional object, which incidentally is of a form that is not faithfully imaged by basic shift-and-add (ϕ3.6.2), is shown in Figure 6.2a. This particular object is seven pixels across and therefore possesses six complex visibility zeros (cf. ϕ6.2). The object is real-valued. Its zero map is therefore mirror-symmetric about the imaginary axis of the complex plane. Figure 6.2b shows the half of the object's zero map lying to the right of the imaginary axis. The pixel spacing Δx has been normalised to unity in Figure 6.2, so that the principal range of u (cf. the final paragraph of ϕ6.2) is the interval $[-0.5;0.5)$. Black represents a pixel value of one and white represents a pixel value of zero in the zero maps displayed in Figure 6.2. The three black dots in Figure 6.2b correspond to three of the six visibility zeros of the object (the other three zeros are mirror images in the jv-axis of the three zeros displayed). The dots have been drawn larger than the actual size of the pixels to make them easier to see. Figure 6.2c shows a one-dimensional computer-simulated speckle image of the object of Figure 6.2a. The speckle image is ideal, because it is both isoplanatic and of finite extent. The speckle image was generated by forming a speckle point spread function as described in Section 4.4, truncating the speckle psf to a sufficiently small extent in image space (35 pixels) for computational convenience, and convolving the truncated psf with the object. The diffraction-limited point spread function of the telescope was modelled as occupying a single pixel in image space, by setting D = SIZE when generating the speckle images (ϕ4.4). Figure 6.2d shows the zero map of the speckle image of Figure 6.2c. As with Figure 6.2b, only the right hand half of the zero map is displayed. Since the speckle image is ideal, $ZS_m(w)$ is the sum (6.22) of ZF(w) and the zero map $ZH_m(w)$ of the speckle transfer function. Three of the numerous black dots in Figure 6.2d therefore originate from ZF(w) and are in exactly the same positions as the three black dots of Figure 6.2b. The remainder of the black dots in Figure 6.2d originate from $ZH_m(w)$ and are positioned randomly. Averaging the zero maps of many statistically independent speckle images yields an averaged zero map

$$\langle ZS_m(w)\rangle_M = \langle ZF(w) + ZH_m(w)\rangle_M = ZF(w) + \langle ZH_m(w)\rangle_M , \qquad (6.23)$$

which is shown for M = 300 speckle images similar to Figure 6.2c in Figure 6.2e. The contribution of ZF(w) to $\langle ZS_m(w)\rangle_M$ is clearly visible in Figure 6.2e, because the nonzero pixels of ZF(w) have values of unity whereas the value of $\langle ZH_m(w)\rangle_M$ is everywhere much less than unity. The contribution from $\langle ZH_m(w)\rangle_M$ is hardly visible in Figure 6.2e. The particular complex zeros of each $S_m(w)$ which originate from the object, and which give rise to the pixels of approximately unit value in $\langle ZS_m(w)\rangle_M$, are termed steadfast zeros because their positions in the zero maps remain steadfast for all m. The positions of the steadfast zeros in the averaged zero map define the complex zeros $\{\omega\}_f$ of the object, from which the object itself may be computed as described in Section 6.2. Figure 6.2f shows an image computed in this manner from the positions of the three black dots in Figure 6.2e and their fellows on the left hand side of the jv-axis. The minor differences between Figure 6.2f and Figure 6.2a are a consequence of the quantisation of w in the computer representation of the zero maps.

Figure 6.2 (facing page): Illustrating the principle of zero-and-add.

(a) A one-dimensional object f(x).

(b) Zero map of (a). The three black dots characterise f(x).

(c) Typical computer simulated ideal speckle image of (a). This image is the convolution of f(x) with a speckle psf $h_m(x)$ of finite extent.

(d) Zero map of (c). The numerous black dots characterise both f(x) and the random $h_m(x)$.

(e) Average of M = 300 statistically independent zero maps similar to (d). Only the pixels characterising f(x) remain prominent after the averaging.

(f) Image computed from the positions of the three black dots in (e).

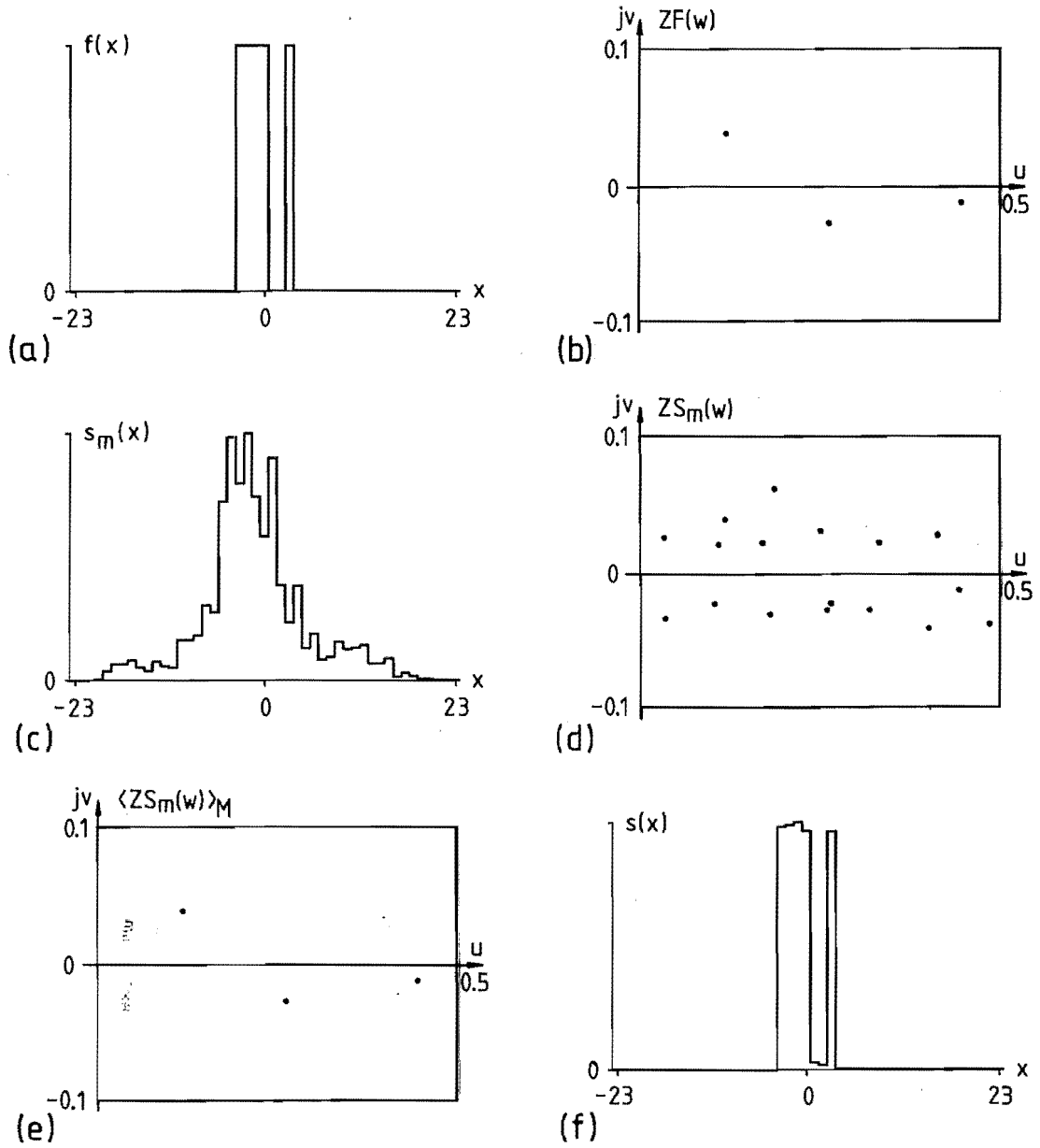


Figure 6.2 (continued):

The speckle images were generated with SIZE = 256, PR = 10, CL = 200, D = SIZE. The speckle psfs were truncated to 35 pixels wide, symmetrical about the origin of image space, before convolution with the object. The pixel spacing Δx (cf. Figure 6.1a) is normalised to unity in the images displayed here and in the following figures of this chapter. The zero maps comprise 129 × 129 pixels here and in Figure 6.3, and 100 × 100 pixels in the remaining figures of this chapter.

6.4 A ZERO-AND-ADD ALGORITHM FOR CONTAMINATED SPECKLE IMAGES

The ideal speckle images used to demonstrate the zero-and-add principle in Section 6.3 are isoplanatic within the limitations of the computer's 32-bit floating-point representation of numbers. Speckle images obtained in practice cannot be expected to be so perfectly isoplanatic. Factors such as quantisation of the speckle images by photon limiting, truncation of the detected speckle images, detector nonlinearities, additive noise, and the inherent nonisoplanatism of short exposure imaging through the Earth's atmosphere, all cause deviations from perfect isoplanatism. The deviations are conveniently represented by a contamination term $c_m(x)$ added to the ideal speckle image (6.19). The contaminated speckle image

$$s_{cm}(x) = f(x) \odot h_m(x) + c_m(x) \tag{6.24}$$

is, almost inevitably, not a convolution of any function with f(x) [cf. Bates et al 1976a ϕ3]. In the visibility domain, $S_{cm}(w)$ is not exactly a product of F(w) with an $H_m(w)$, so the complex zeros of F(w) do not necessarily appear in $ZS_{cm}(w)$. However, an infinitesimally small change to $s_m(x)$ causes an infinitesimally small change to the complex visibility zeros of $s_m(x)$. One can therefore imagine the complex zeros in $ZS_m(w)$ being displaced along continuous loci from their original positions as $s_m(x)$ is gradually contaminated. The complex zeros of F(w), which are present in their proper positions in $ZS_m(w)$, can be considered to be also present in $ZS_{cm}(w)$ but displaced from their proper positions by the contamination. Provided the contamination is not too severe, the displacements are small and so these complex zeros are almost steadfast in $ZS_{cm}(w)$. They are then referred to as quasi-steadfast zeros of $S_{cm}(w)$. The quasi-steadfast zeros give rise to clusters of points in $\langle ZS_{cm}(w)\rangle_M$ which are recognisably denser than the background of points in $\langle ZS_{cm}(w)\rangle_M$ due to the $H_m(w)$. The positions of these clusters constitute an estimate of the values of the complex zeros of F(w). They allow an image of the object to be formed by zero-and-add when the speckle images are contaminated.

The quasi-steadfast zeros of $S_{cm}(w)$ may be made recognisable by convolving the averaged zero map $\langle ZS_{cm}(w)\rangle_M$ with a blurring function

blur(w), whose effective diameter should be comparable to the extents of the clusters of points in $\langle ZS_{cm}(w)\rangle_M$ representing the quasi-steadfast zeros. The blurring is equivalent to representing the complex visibility zeros of each $s_{cm}(x)$ by discs shaped like blur(w) instead of by points in $ZS_{cm}(w)$. The closely spaced discs from the quasi-steadfast zeros overlap to produce recognisable local maxima in $\langle ZS_{cm}(w)\rangle_M$.

The random background of points from $H_m(w)$ in the averaged zero map may happen to cluster more densely in some regions of the complex plans than in others, due to the statistical properties of the seeing. These denser regions of the background are termed <u>obdurate zeros</u>, since they appear similar to quasi-steadfast zeros in the averaged zero map but originate from the seeing instead of the object. When the averaged zero map is blurred by convolving it with blur(w), the obdurate zeros create local maxima additional to the local maxima from the quasi-steadfast zeros. The quasi-steadfast zeros generate local maxima of considerably greater value than do the obdurate zeros when the contamination is mild. However, as the severity of the contamination increases, the values of the local maxima due to the quasi-steadfast zeros become progressively less, eventually becoming comparable to the values of the local maxima due to the obdurate zeros. The quasi-steadfast zeros are then difficult to distinguish from the obdurate zeros in the blurred averaged zero map.

The obdurate zeros in $\langle ZS_{cm}(w)\rangle_M$ can be countered with the aid of speckle images $s_{com}(x)$ of an unresolved reference object viewed through statistically similar seeing to that applying to the $s_{cm}(x)$. This is reminiscent of the use of observations of an unresolved object to characterise and compensate for the effects of the seeing in many other speckle processing techniques (cf. Chapter 3). The visibility of an unresolved object $f_o(x) \approx \delta(x)$ is essentially constant, so the averaged zero map $\langle ZS_{com}(w)\rangle_{Mo}$ for the unresolved object contains no quasi-steadfast zeros. It does however contain the obdurate zeros due to the seeing. Subtracting $\langle ZS_{com}(w)\rangle_{Mo}$ from $\langle ZS_{cm}(w)\rangle_M$, and blurring the difference by convolving it with the blur(w) mentioned above, produces the <u>ZAA map</u>

$$ZAS(w) = \left[\langle ZS_{cm}(w)\rangle_M - \langle ZS_{com}(w)\rangle_{Mo}\right] \odot \text{blur}(w).$$

The quasi-steadfast zeros are largely unaltered by the subtraction but the obdurate zeros are largely suppressed. This makes the quasi-steadfast zeros recognisable in ZAS(w) in the presence of greater contamination than when $\{s_{com}(x)\}$ is not invoked.

The considerations outlined above lead to the following practical algorithm for zero-and-add with contaminated speckle images:

(a) Obtain ensembles of one-dimensional speckle images of an object of interest and of an unresolved reference object viewed through statistically similar seeing. Call these ensembles $\{s_{cm}(x)\}$ and $\{s_{com}(x)\}$ respectively (the subscript c is explicit recognition that the speckle images are inevitably contaminated in practice).

(b) Form averaged zero maps (ϕ6.3) $\langle ZS_{cm}(w)\rangle_M$ for the object and $\langle ZS_{com}(w)\rangle_{Mo}$ for the reference object. Compute the difference zero map

$$\langle ZS_{cm}(w)\rangle_M - \langle ZS_{com}(w)\rangle_{Mo}. \tag{6.25}$$

(c) Estimate the extent E of the object in image space. From the estimate E_e of E compute

$$N_e = E_e/\Delta x - 1, \tag{6.26}$$

where Δx is the spacing of the pixels of image space. N_e is an estimate of the number N of complex visibility zeros the object possesses (ϕ6.2), and hence of the number of quasi-steadfast zeros to be identified in the ZAA map in step (g). Since the object f(x) is real and non-negative in astronomy (cf. ϕ2.2), its extent is necessarily half that of its autocorrelation [Bates and Fright 1984 ϕ6]. E_e may therefore be taken as half of the effective extent of the estimated autocorrelation of f(x) (3.21, 3.22) produced from $\{s_{cm}(x)\}$ and $\{s_{com}(x)\}$ by speckle interferometry. A priori knowledge of the object may also be helpful in establishing E_e. It is better for E_e to be larger than E rather than smaller if its value is uncertain.

(d) Define a blurring function blur(w), whose value is unity at the

origin and whose effective diameter is appreciably less than $1/N_e$. $1/N_e$ is the expected average horizontal spacing in the complex plane of the complex zeros of the object. A suitable blurring function is the Gaussian function

$$\text{blur}(w) = e^{-r|w|^2}, \tag{6.27}$$

where r is a real constant determining the diameter of blur(w).

(e) Compute the ZAA map

$$\text{ZAS}(w) = \left[\langle \text{ZS}_{cm}(w)\rangle_M - \langle \text{ZS}_{com}(w)\rangle_{Mo}\right] \circledcirc \text{blur}(w). \tag{6.28}$$

(f) Define a threshold μ which represents an estimate of the maximum value of the background of detail in ZAS(w) other than quasi-steadfast zeros. A suitable value for the threshold might be

$$\mu = 0.2 . \tag{6.29}$$

(g) Locate all of the local maxima of ZAS(w) whose values are greater than μ. If there are more than N_e such local maxima, choose from them the N_e maxima with the greatest values, otherwise choose all of them. Let N' equal the number of local maxima actually chosen.

(h) Form a set $\{\omega_p\}$ of complex numbers equal to the positions in the ZAA map of the local maxima chosen in step (g). Compute the values $\{s_n\}$ from $\{\omega_p\}$ as described in Section 6.2, using (6.16), (6.15), and (6.14). From $\{s_n\}$ form the zero-and-add image

$$s_{ZA}(x) = \sum_{n=0}^{N'} s_n \delta(x - (n - N'/2)\Delta x). \tag{6.30}$$

The zero-and-add algorithm for contaminated speckle images is illustrated in Figure 6.3. In this example, ideal speckle images similar to Figure 6.2c were stored as arrays of eight bit integers. The quantisation of value associated with eight-bit representation contaminated the speckle images in a comparatively mild manner.

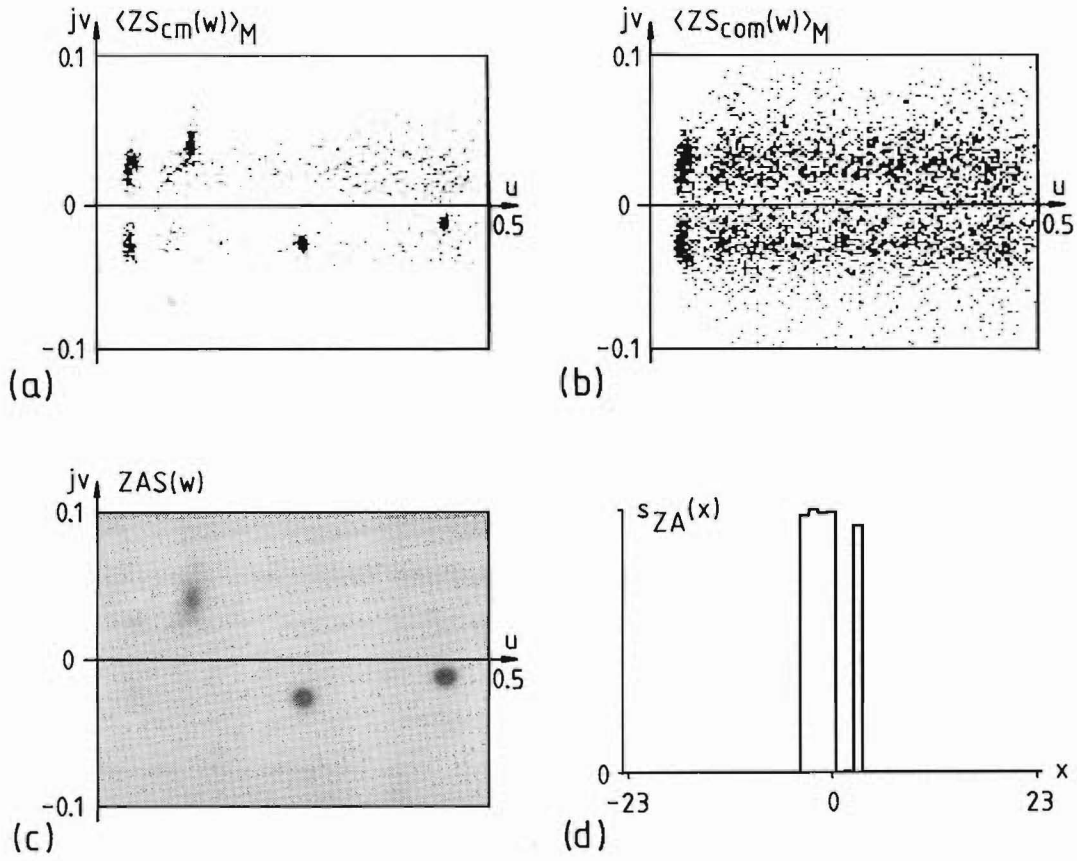


Figure 6.3: Illustrating zero-and-add with contaminated speckle images.

(a) Average of zero maps of M = 300 statistically independent contaminated speckle images of the object shown in Figure 6.2a. The speckle images, initially similar to Figure 6.2c, were contaminated by being quantised to 256 levels.

(b) Average of zero maps of 300 statistically independent speckle images of a point object. The speckle images were generated and contaminated as for (a).

(c) ZAA map (6.28), or blurred difference between (a) and (b), of the object in Figure 6.2a.

(d) Zero-and-add image computed from the positions of the maxima of the prominent dark regions in (c).

Figure 6.3a shows the averaged zero map formed from 300 contaminated speckle images of the object shown in Figure 6.2a. The complex zeros of the object are quasi-steadfast in the $S_{cm}(w)$ because of the contamination, and appear as clusters of dark pixels in Figure 6.3a. Obdurate zeros are visible in Figure 6.3a as the two footprint-like dark regions above and below the u-axis close to the jv-axis. Figure 6.3b shows the averaged zero map formed from 300 contaminated speckle images of a point object. The two footprints of Figure 6.3a are clearly visible in Figure 6.3b, together with other detail of the random background of complex zeros resulting from the seeing. Note that the displays of Figures 6.3a-c have been normalised so that the most positive pixel and the most negative pixel in each zero map appears as black and white respectively. Figure 6.3c shows the ZAA map computed from Figures 6.3a and b according to (6.28), using a Gaussian blurring function (6.27) whose half-intensity diameter is equal to about 8% of the displayed widths of the zero maps. The quasi-steadfast zeros are prominent in Figure 6.3c and the obdurate zeros (the footprints in Figures 6.3a) have been suppressed. Figure 6.3d shows the zero-and-add image (6.30), computed from Figure 6.3c, which closely resembles the object (Figure 6.2a).

In the example just presented the contamination is mild and causes no appreciable degradation of the zero-and-add image. The effects of more severe contamination are considered in the following section.

6.5 EFFECTS OF CONTAMINATION ON ZERO-AND-ADD

This section presents computational studies of the effects on zero-and-add of three particular types of contamination of speckle images: pseudo-random additive noise, photon limiting, and truncation of the speckle images. For each of the computational examples presented in this section, ideal speckle images with statistical properties identical to those of Section 6.3 were generated and then contaminated as described in the individual subsections below.

The zero maps shown in the figures of this section are each ZAA maps (ϕ6.4) depicted as contour plots. Each ZAA map has six

contours linearly spaced across the top three quarters of the range from zero to the maximum positive value in the map. The contour plots have an implicit threshold because of the positive value of the lowest contour. The negative parts and nondescript background of the ZAA maps are therefore not displayed. The diameter of the blurring function (6.27) is about 10% of the displayed width of the zero maps. Each contour is marked with a small dot indicating the positive or "uphill" side of the contour.

6.5.1 ADDITIVE NOISE

Additive noise in speckle images can arise in practice from detector noise and from "sky background" radiation. In the examples of this subsection, additive noise was introduced by adding a pseudo-random number to each pixel of ideal speckle images like the one in Figure 6.2c. The pseudo-random numbers were uncorrelated, and were uniformly distributed between zero and a certain positive maximum value unless otherwise noted below. The severity of the contamination is expressed quantitatively by the relative contamination c_r, defined by

$$c_r = \frac{\sqrt{\int_{T_c} c_m^2(x)\ .dx}}{\sqrt{\int_{T_c} s_m^2(x)\ .dx}}, \tag{6.31}$$

This is the ratio of the RMS value of the noise to the RMS value of the ideal speckle image over the support T_c of the contaminated speckle images. T_c contains 45 pixels, comprising the support of the ideal speckle images plus a few extra pixels. The quality of reconstruction of the object is expressed quantitatively by the reconstruction error e_r, defined by

$$e_r = \frac{\sqrt{\int_{T_z} \left(c\, s_m(x) - f(x)\right)^2\ .dx}}{\sqrt{\int_{T_z} f\ (x)\ .dx}} \tag{6.32}$$

where T_z is the estimated support of the object and the scaling constant c is chosen to minimise e_r. This definition of e_r matches the definition of c_r but otherwise has no particular physical significance.

Figure 6.4 shows the effect of additive noise on the zero-and-add reconstruction of the object of Figure 6.2. The left side of Figures 6.4a-d are ZAA maps formed with increasing amounts of contamination. Since the zero maps are symmetrical about the jv-axis, only half of the principal range of u is displayed. The images reconstructed from the zero maps (cf. φ6.4) are shown on the right of Figures 6.4a-d. The increasing level of additive noise has two obvious effects on the ZAA maps: the quasi-steadfast zero clusters each become spread out, and they each migrate towards the u-axis. (The way to judge the width of each cluster from the contour plots is to take the diameter of, say, the second contour down from the maximum contour in each cluster.)

The migration of the clusters towards the u-axis impairs the fidelity of the zero-and-add images. The loci of the maxima of the clusters vary only a little in u as jv reduces due to contamination. Figure 6.4f shows the image obtained from complex zeros lying along the real axis, with real parts equal to the real parts of the steadfast zeros in Figure 6.4a. This represents a crude extrapolation of the migration to its limit. The image so obtained is symmetrical and is quite unlike the object. Figure 6.4e plots the reconstruction error against the relative contamination in Figures 6.4a-d. The reconstruction error increases rapidly with c_r when c_r is small, because of the migration of the quasi-steadfast zero clusters.

In the ZAA maps of Figure 6.4 it is apparent that, for this object, the spreading out and migration of the quasi-steadfast zero clusters are more pronounced for clusters closer to the jv-axis and further from the u-axis. An heuristic explanation for the differing susceptibility to contamination of different clusters in a zero map can be made by considering a geometrical picture of visibility magnitudes. Imagine that the magnitude $|F(w)|$ of the visibility of the object f(x), shown as $|S(w)|$ in Figure 6.1d, is a landscape punctuated by wells (the zeros) leading down to negative infinity,

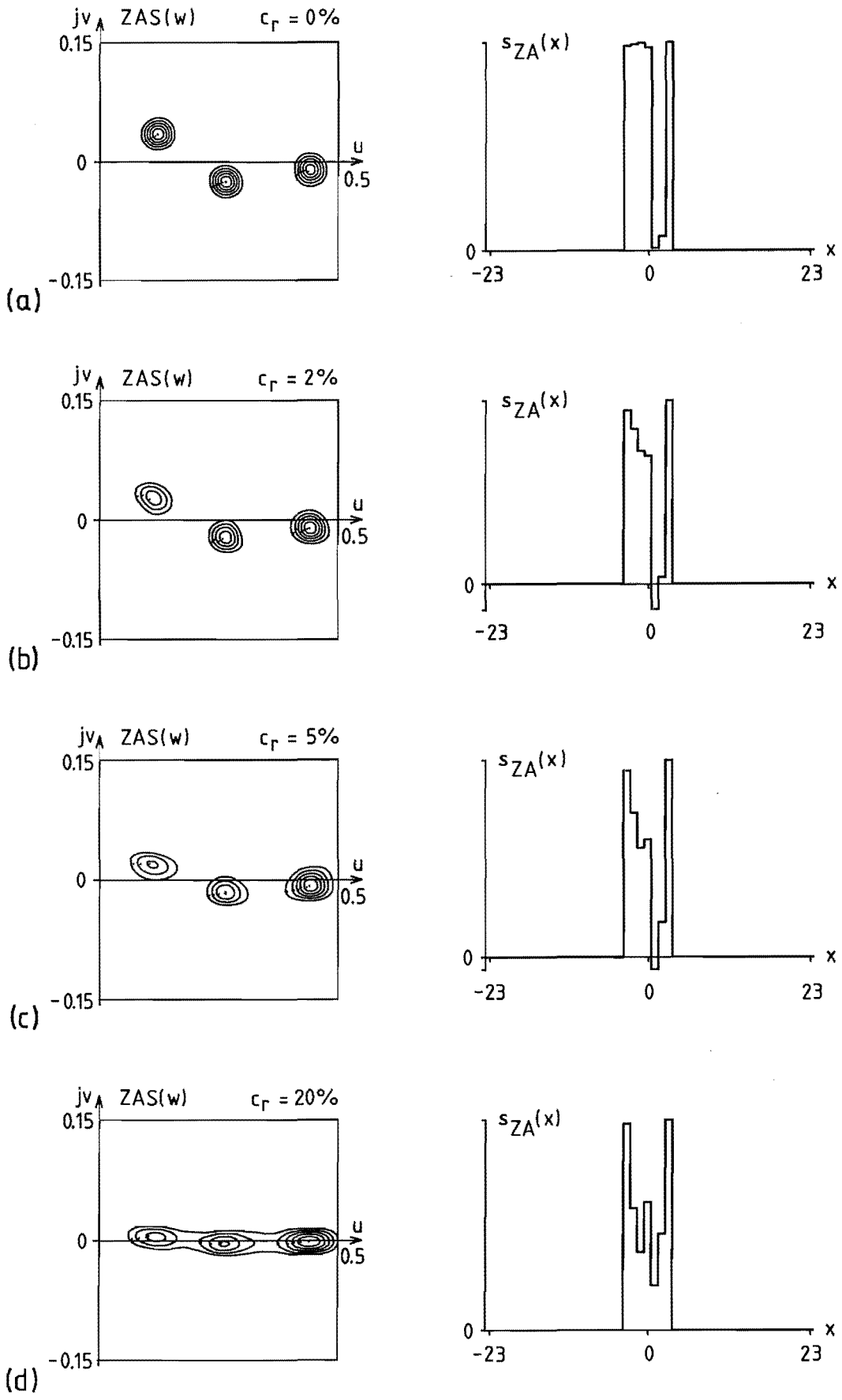

jv
ZAS(w)
c_r = 0%
0.15
0
-0.15
u
0.5
(a)
$s_{ZA}(x)$
0
x
-23
0
23
jv
ZAS(w)
c_r = 2%
0.15
0
-0.15
u
0.5
(b)
$s_{ZA}(x)$
0
x
-23
0
23
jv
ZAS(w)
c_r = 5%
0.15
0
-0.15
u
0.5
(c)
$s_{ZA}(x)$
0
x
-23
0
23
jv
ZAS(w)
c_r = 20%
0.15
0
-0.15
u
0.5
(d)
$s_{ZA}(x)$
0
x
-23
0
23

Figure 6.4

and with hills rising between the wells. The wells originate from destructive interference between the Fourier components from the different pixels of the object at certain places in the complex plane. There are comparatively few wells for an object with few pixels. The wells in general need not be of equal sharpnesses or widths - some of the interference nulls in the spectrum of the object may be sharper than others. Multiplying F(w) by $H_m(w)$ punctures the landscape with a large number of new wells in random positions. The shape of the landscape around the original wells is altered by the presence of the new wells. However the sharpest of the original wells should remain, on average over many m, the sharpest original wells after the puncturing. When contamination $C_m(w)$ is added to the visibility, it

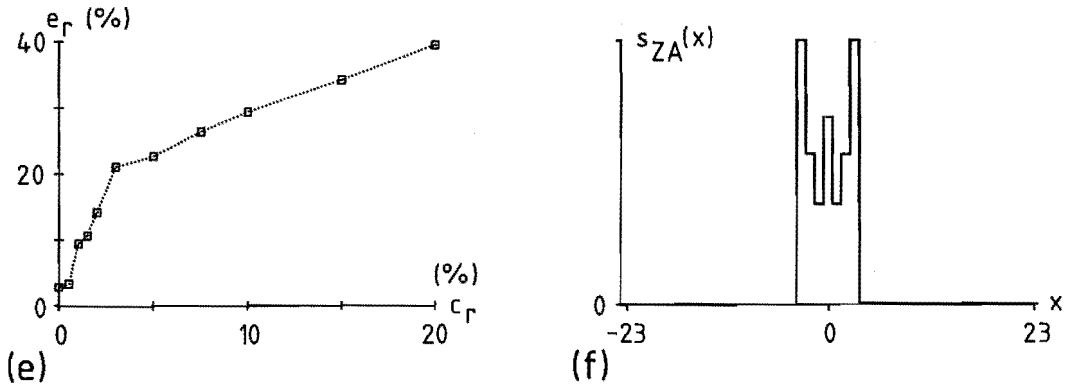


Figure 6.4 (facing page and above): The effect of additive noise on zero-and-add reconstruction of the object in Figure 6.2a.

(a) - (d) ZAA maps (left) and the corresponding zero-and-add images (right), each formed from M = 300 statistically independent speckle images similar to Figure 6.2c contaminated by pseudo-random additive noise (see text) in the amounts shown on the figures.

(e) Dependence of reconstruction error (6.32) on relative contamination (6.31) for this example.

(f) Hypothetical image obtained when the migration of the quasi-steadfast zero clusters toward the u-axis, apparent in (a) - (d), is extrapolated all the way to the u-axis, along lines parallel to the jv-axis from the positions of the steadfast zeros in (a).

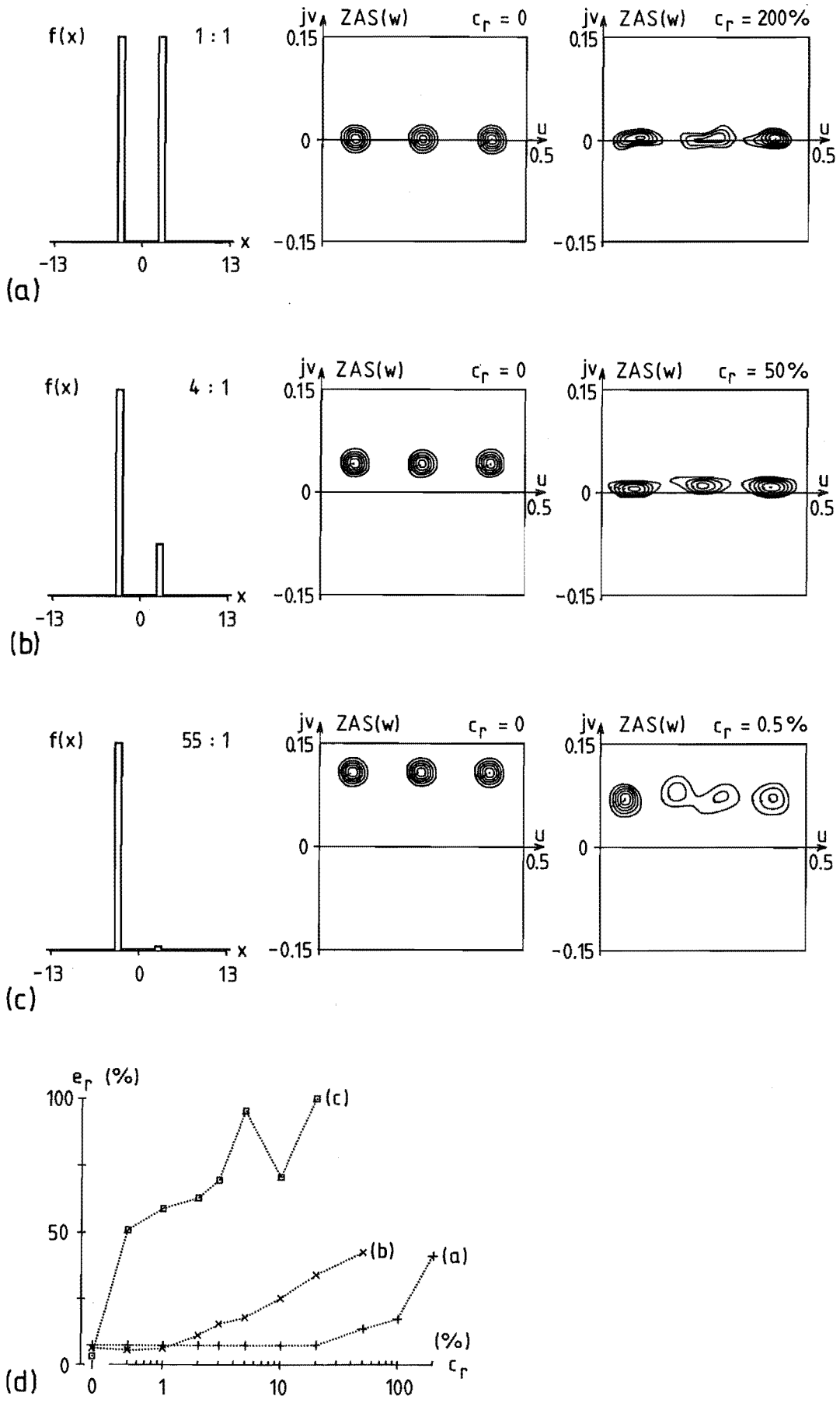

f(x)
1 : 1
jv ZAS(w)
c_r = 0
jv ZAS(w)
c_r = 200%
0.15
0
-0.15
0.5
u
x
-13
0
13
(a)
f(x)
4 : 1
jv ZAS(w)
c_r = 0
jv ZAS(w)
c_r = 50%
(b)
f(x)
55 : 1
jv ZAS(w)
c_r = 0
jv ZAS(w)
c_r = 0.5%
(c)
e_r (%)
100
50
0
(c)
(b)
(a)
(%)
c_r
0
1
10
100
(d)

Figure 6.5

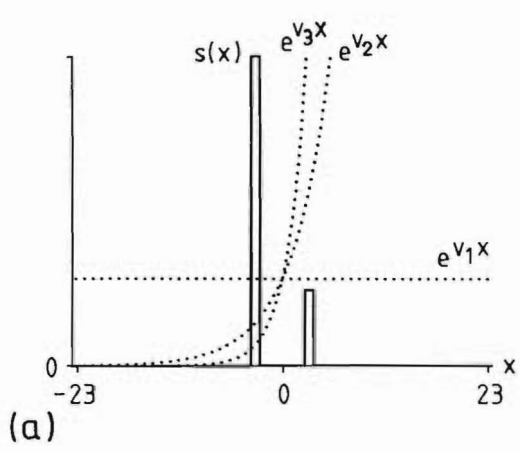
interferes with $F(w)H_m(w)$. Provided $|C_m(w)|$ is small compared to $|F(w)H_m(w)|$ on average over m, the presence of $C_m(w)$ affects the landscape $|S_m(w)|$ by shifting the wells slightly. The wells which are sharpest and surrounded by the steepest hills can be expected by be shifted the least. This suggests that the quasi-steadfast zero clusters associated with the strongest fringes in the object's visibility F(w) should be the zeros least affected by additive noise or by other forms of contamination.

Figure 6.5 shows ZAA maps of three different binary objects, formed from speckle images with and without contamination by additive noise. The binary objects, shown on the left of Figures 6.5a-c, each span seven pixels and have unresolved parts with relative strengths of 1:1, 4:1, and 55:1 respectively. The changing intensity ratio is accompanied by a displacement in the jv-direction of the steadfast zeros in ZAA maps formed from ideal speckle images (centre of Figures 6.5a-c). The right side of Figures 6.5a-c shows the ZAA maps obtained from speckle images contaminated to a degree sufficient to make the reconstruction error just exceed 40%. The variation of e_r with c_r for the three objects is shown in Figures 6.5d. e_r is greater than zero when c_r is zero because of the quantisation of w in the computer

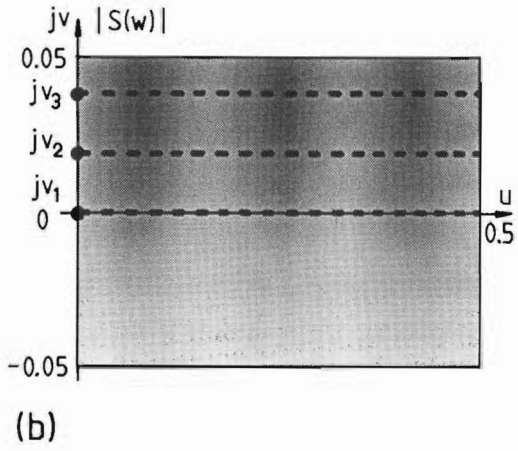
Figure 6.5 (facing page): Demonstrating the dependence of a steadfast zero's susceptibility to contamination on its position in the complex plane.

(a) - (c) (each) Left - Binary object with unresolved parts.
Centre - ZAA map from 300 statistically independent ideal speckle images of the object.
Right - ZAA map from 300 statistically independent speckle images of the object contaminated, by additive noise as in Figure 6.4, by the amounts shown here.

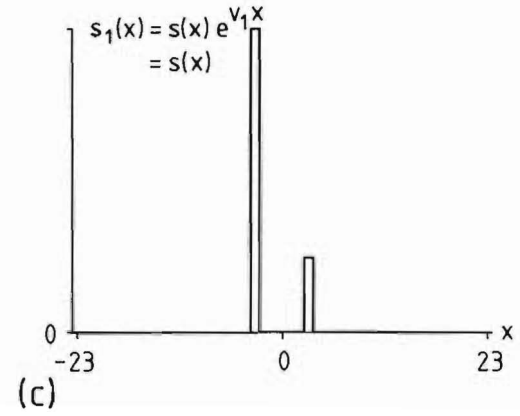
(d) Dependence of reconstruction error on relative contamination for the three objects in (a) - (c).



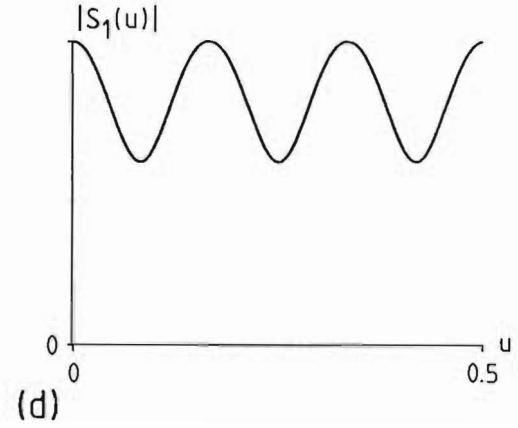
(a)



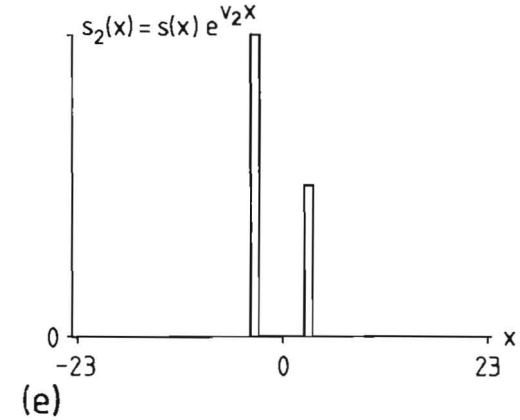
(b)



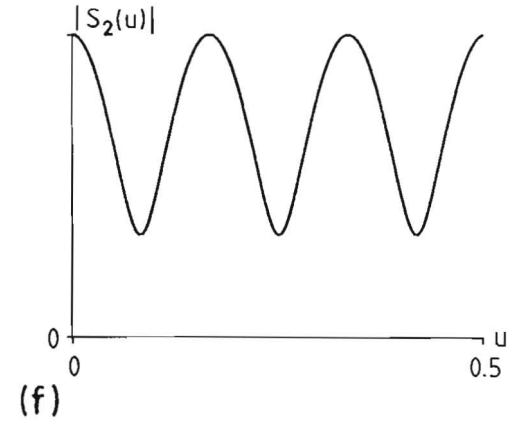
(c)



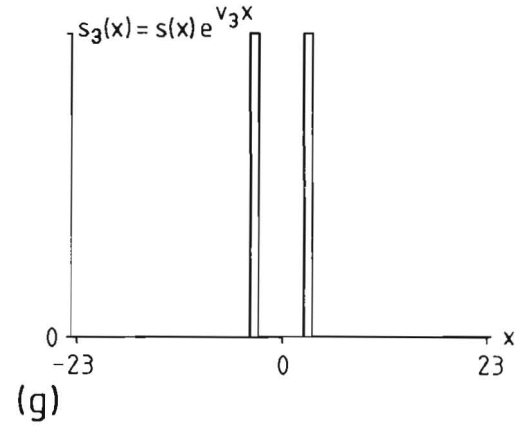
(d)



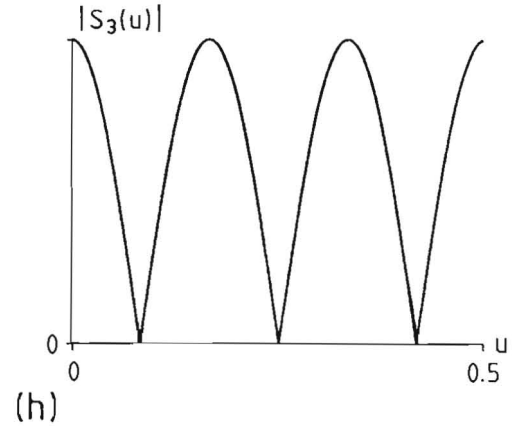
(e)



(f)



(g)



(h)

Figure 6.6

representation of the zero maps. The ZAA map and the zero-and-add image of the object with 1:1 strength ratio are little affected by additive noise for c_r up to 20%, since the object's complex visibility zeros lie on the u-axis and therefore the quasi-steadfast zeros in the ZAA map cannot migrate towards the u-axis. The quasi-steadfast zeros eventually migrate along the u-axis when c_r becomes large enough (Figure 6.5a right). The quasi-steadfast zeros in the ZAA maps of the 4:1 and 55:1 objects each migrate towards the jv-axis as c_r increases. The migration occurs much more rapidly with increasing c_r for the 55:1 object than for the 4:1 object. The quasi-steadfast zeros of the 55:1 object are not identifiable in the ZAA map without a priori knowledge of the object for relative contaminations of 0.5% and above.

The example of Figure 6.5 shows clearly that the zero-and-add image of a binary object with unresolved parts becomes more sensitive to additive noise as the ratio of the strengths of the two parts increases. This is a consequence of the relative magnitudes of $F(w)H_m(w)$ and $S_m(w)$ being different at different positions in the complex plane [cf. Bates et al 1976a φ2]. In the expression (6.3) for the visibility of an image, w may be expanded into its real and

Figure 6.6 (facing page): The significance of slices through S(w) at constant v.

(a) An image s(x), and three real exponential functions $e^{2\pi vx}$ for $v = v_1, v_2, v_3$.

(b) Visibility S(w) of s(x), and three horizontal lines through S(w) at $v = v_1, v_2, v_3$.

(c), (e), (g) Three images equal to the product of s(x) with the three exponential functions in (a).

(d), (f), (h) The visibilities of the images (c), (e), (g), which are equal to the value of S(w) along the corresponding horizontal lines in (b).

imaginary parts to give

$$S(w) = \int_{-\infty}^{\infty} s(x)\, e^{-j2\pi wx}\, .dx$$

$$= \int_{-\infty}^{\infty} s(x)\, e^{-j2\pi(u + jv)x}\, .dx$$

$$= \int_{-\infty}^{\infty} s(x)\, e^{2\pi vx}\, e^{-j2\pi ux}\, .dx$$

$$= \mathbf{F}[s(x)\, e^{2\pi vx}]. \qquad (6.33)$$

Thus, a slice through S(w) parallel to the u-axis at a certain value of v gives the conventional Fourier transform, as a function of u, of an image equal to s(x) multiplied by the real exponential function $e^{2\pi vx}$ (see Figure 6.6). When $|v|$ is large, the product $s(x)e^{2\pi vx}$ is dominated by the value of s(x) at one or the other edge of its support. The expected value of a speckle image, which is equal to the long exposure image, is large at the centre of the speckle image and diminishes away from the centre. The ratio of contamination $c_m(x)$ with uniform average magnitude throughout the contaminated speckle image $s_{cm}(x)$ to the ideal component $f(x) \circledcirc h_m(x)$ is least in the centre of the speckle image and increases away from the centre (see Figure 6.7). Multiplying $s_{cm}(x)$ by the exponential $e^{2\pi vx}$ accentuates the noisy regions at the edge, to a greater degree for greater $|v|$. The expected value of $|C_m(w)|$ compared to the expected value of $|F(w)H_m(w)|$ in the visibility domain is therefore greater for greater $|v|$. Thus, aside from possible intrinsic differences in the sharpnesses of an object's complex visibility zeros as mentioned earlier in this subsection, quasi-steadfast zero clusters can be expected to be less durable in the presence of additive contamination the further they are from the u-axis. The examples of Figure 6.5 illustrate this clearly since the visibilities S(w) of the three objects are identical apart from translations in the jv-direction.

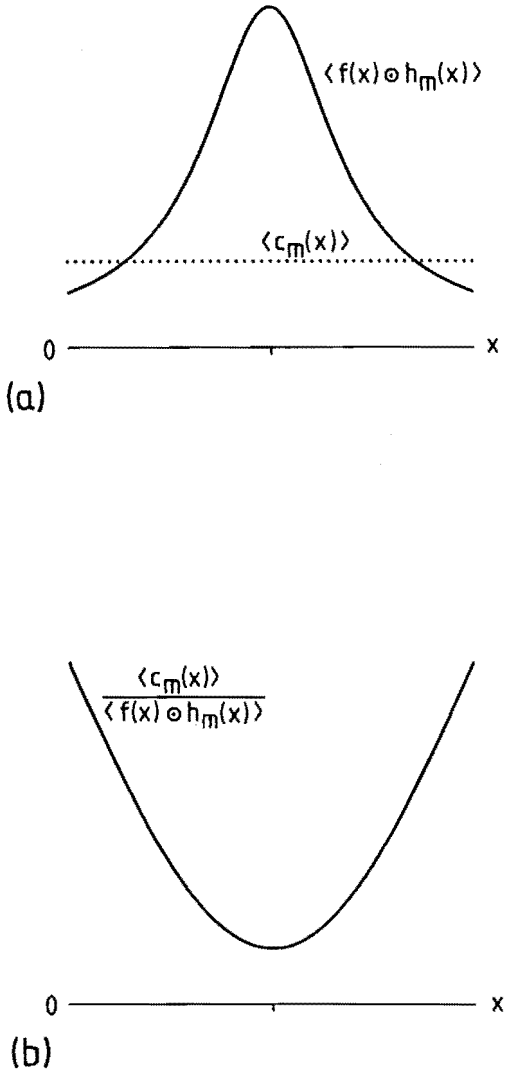


Figure 6.7: The ratio of the expected value of uniformly distributed additive contamination to the expected value of an ideal speckle image (a) is highest at the edges of the speckle image (b).

6.5.2 PHOTON LIMITING

As stated in the final paragraph of Section 6.2, the visibility S(w) of a sampled image s(x) is periodic in the u-direction, with a period of 1 if the sample spacing is normalised to unity. The periodicity permits a principal range of u of [0;1) to be used instead of [-0.5;0.5) in displaying ZAA maps. A ZAA map displayed over the range u = [-0.5;0.5) (or the range u = [0;0.5], cf. ϕ6.3) is somewhat unclear when there is much significant detail near u = ±0.5, because this detail lies near the edges of the display. The principal range of [0;1) is advantageous in that it places u = ±0.5 in the centre of the display. Quasi-steadfast zeros near u = ±0.5 are thereby made easier to see. For this reason, ZAA maps are displayed on u = [0;1) in this subsection instead of u = [0;0.5] as elsewhere.

The effect of photon limiting on zero-and-add is demonstrated in Figure 6.8. The object in this example (Figure 6.8a) is similar to the object of Figure 6.2, but with the wide part being three pixels wide instead of four. Ideal speckle images formed with speckle point spread functions statistically similar to the speckle psf for the speckle image shown in Figure 6.2c were photon limited, with $\bar{N}_p$ photons per speckle image on average, as described in Section 4.6. Pixels near the centre of the photon-limited speckle images received approximately $\bar{N}_p/17$ photons on average each, while pixels near the edges of the speckle images received approximately $\bar{N}_p/600$ photons on average each. Zero-and-add was performed on the photon-limited speckle images yielding the ZAA maps shown in Figures 6.8b-e.

The quasi-steadfast zero clusters in the ZAA maps of Figure 6.8 spread out and migrate towards the u-axis as $\bar{N}_p$ decreases. These effects are similar to the effects of additive noise on ZAA maps (ϕ6.5.1). Figure 6.8f shows the variation with $\bar{N}_p$ of the reconstruction error defined by (6.32). The reconstruction error evidently rises steadily as $\bar{N}_p$ decreases.

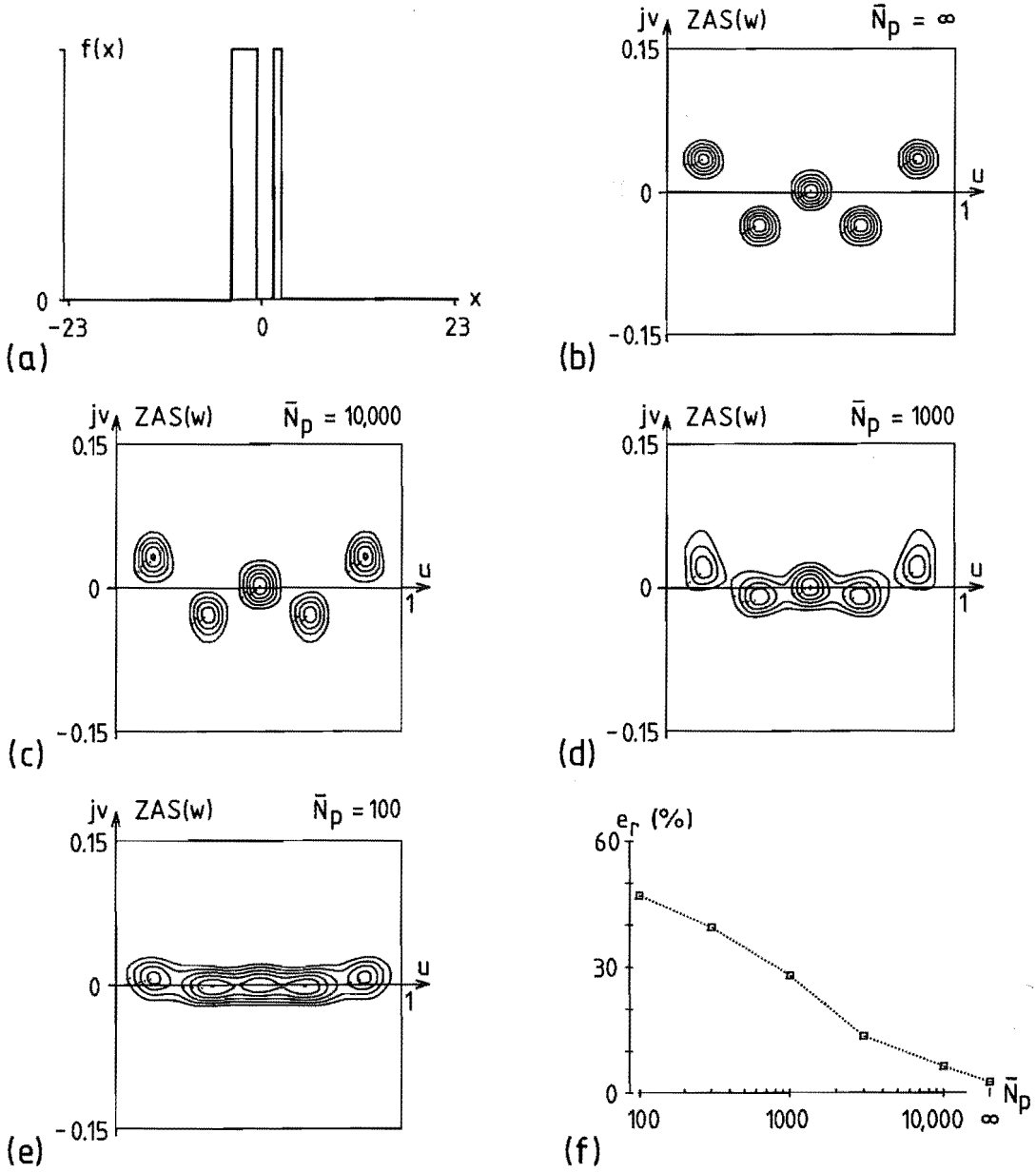


Figure 6.8: The effect of photon limiting of speckle images on zero-and-add reconstruction.

(a) Object.

(b) - (e) ZAA maps formed from 300 statistically independent photon-limited speckle images of (a) with $\bar{N}_p = \infty$ (i.e. no photon limiting) (b), 10,000 (c), 1000 (d), 100 (e).

(f) Dependence of reconstruction error on $\bar{N}_p$ for this example.

The photon-limited speckle images were formed (φ4.6) from ideal speckle images similar to Figure 6.2c.

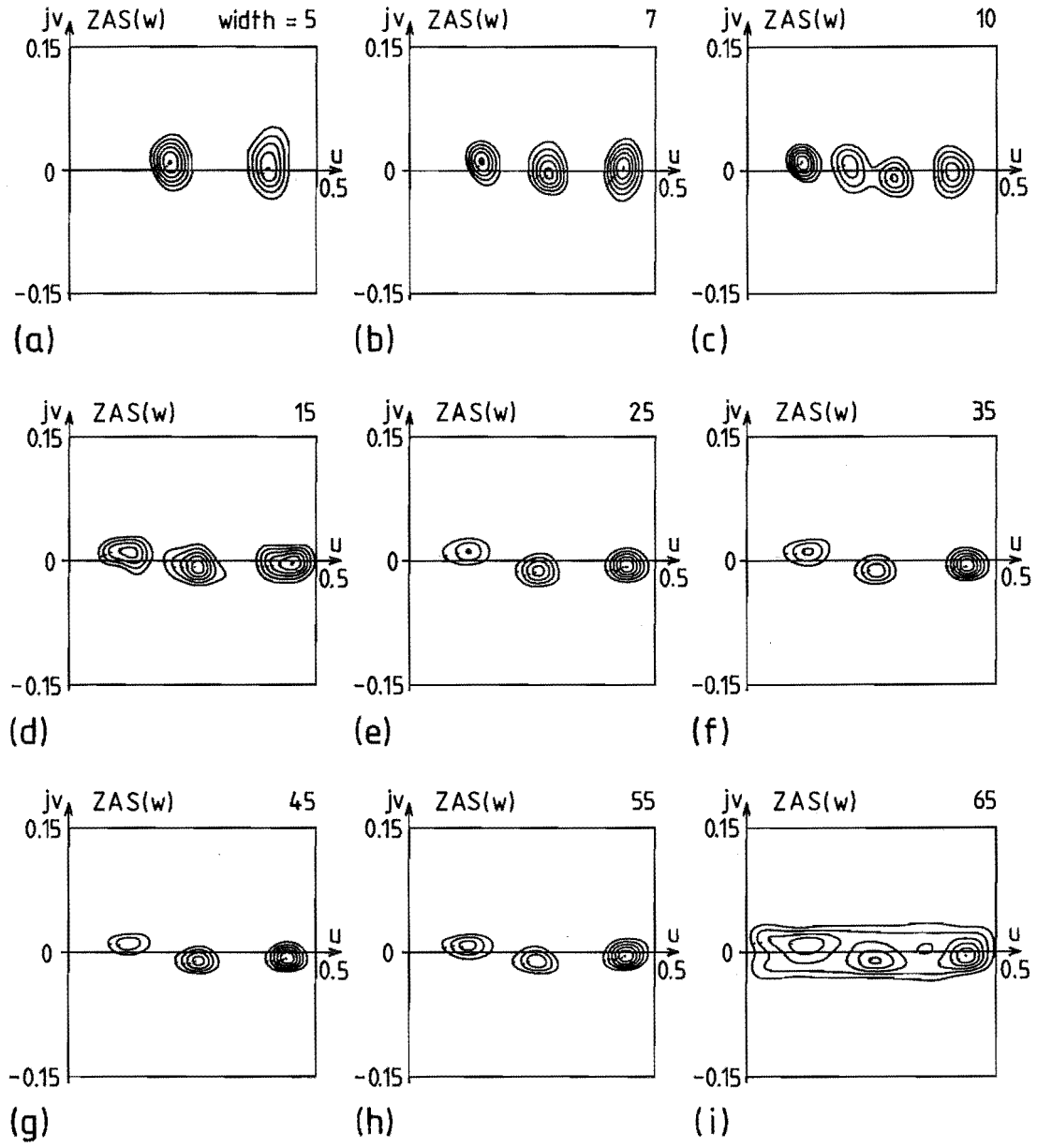


Figure 6.9: The effect on zero-and-add of truncating speckle images with rectangular windows.

(a) - (i) ZAA maps formed from 300 statistically independent truncated speckle images of Figure 6.2a.

The speckle images for each figure were generated by multiplying ideal speckle images, similar to Figure 6.2c except that the speckle psfs were not truncated before convolution with the object, by a rectangular window symmetrically disposed about the origin of image space and with the width shown in the figure.

6.5.3 TRUNCATED SPECKLE IMAGES

Speckle images may be inadvertently truncated during their measurement if they overlap the edges of the detector, or they may be deliberately truncated after measurement if their extents encompass too many pixels for computational convenience. In either case, truncating an otherwise ideal speckle image introduces contamination, because a convolution of one function with another does not in general remain a convolution after it is truncated [cf. Bates et al 1976a ϕ3]. Consequently, zero-and-add does not produce steadfast zeros from truncated speckle images.

For the examples presented in this subsection, ideal speckle images $s_m(x)$ of the object shown in Figure 6.2a were generated in the same way and with the same statistical parameters as the ideal speckle image shown in Figure 6.2c, except that here the speckle point spread functions were not truncated before being convolved with the object. The ideal speckle images were truncated after the convolution by multiplying them with a window function, and were then processed by zero-and-add (ϕ6.4).

Figure 6.9 shows the ZAA maps obtained for rectangular window functions of various widths. The ZAA maps for the narrowest windows, 10 or fewer pixels wide, bear little resemblance to the zero map of the object (cf. Figure 6.4a). The ZAA maps of the wider windows, 15 or more pixels wide, show the appropriate number of quasi-steadfast zero clusters for the extent (seven pixels) of the object, but the clusters are broader and closer to the u-axis than are the steadfast zeros of the uncontaminated ZAA map (Figure 6.4a). The positions of the local maxima of the clusters, and hence the reconstruction errors (6.32) of the zero-and-add images, are similar for all of the wider rectangular windows. Figure 6.10 shows the widths of the windows compared to the expected value of the unwindowed speckle images, the ratios of the expected areas of the windowed speckle images to the unwindowed speckle images, and the reconstruction errors of the zero-and-add images formed from the ZAA maps of Figure 6.9 against window width. The reconstruction error does not decrease uniformly with increasing window width as one might expect. Instead it remains

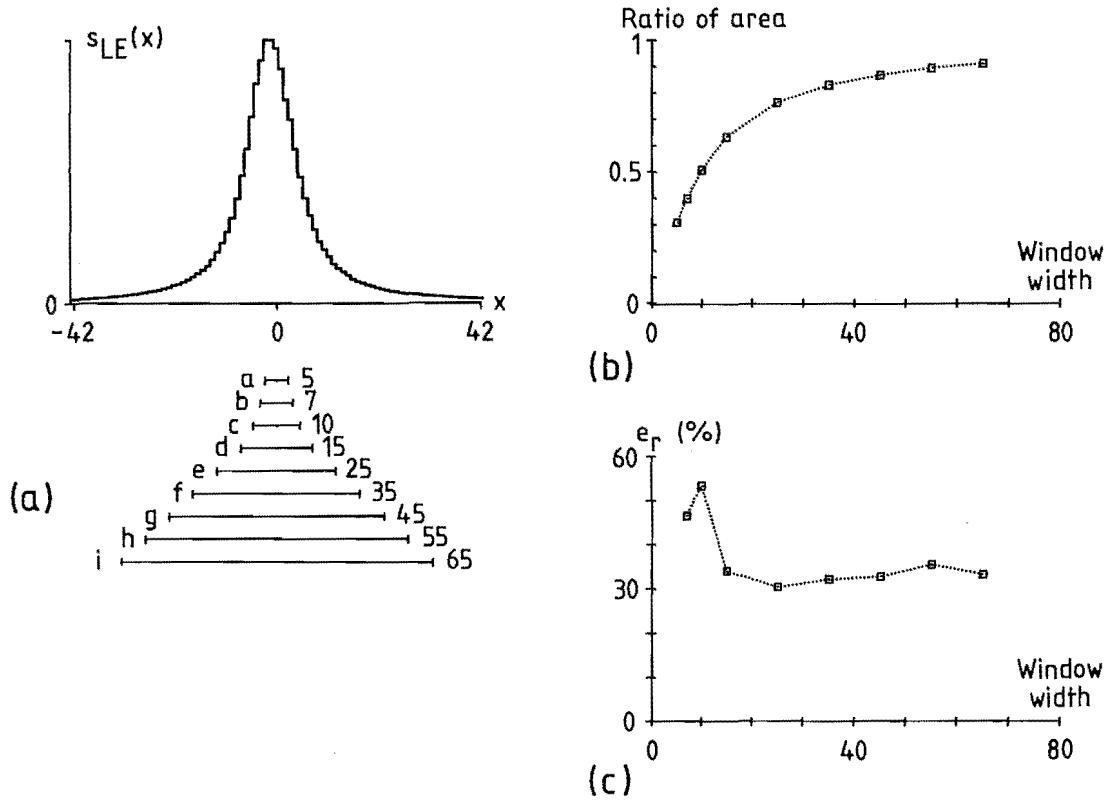


Figure 6.10: Truncation of speckle images by rectangular windows (continued from Figure 6.9).

(a) Illustrating the widths of the windows of Figures 6.9a-i in comparison to the long exposure image of the object.

(b) Ratio of the expected area of the windowed speckle images to the unwindowed speckle images, indicating the average proportion of the "energy" of the ideal speckle images passed by each window.

(c) Dependence of reconstruction error on window width in Figure 6.9.

roughly equal and comparatively high (cf. Figures 6.4i, 6.8f) for all of the windows 15 or more pixels wide. This remarkable result is not caused by numerical instability of the zero-finding algorithm at the large window widths, since zero-and-add performed on ideal speckle images with extents of 65 pixels produced perfectly steadfast zeros in their correct positions in the ZAA map.

Figure 6.11 shows zero-and-add reconstructions of the object of Figure 6.2a from ideal speckle images truncated by windows of various shapes, as shown on the left of Figures 6.11a-c. The ideal speckle images were statistically identical to the ones generated for the example of Figure 6.9. The three windows are each 45 pixels wide. The window of Figure 6.11a is identical to the rectangular window of Figure 6.9g. The window of Figure 6.11b was formed by taking a rectangular function 3 pixels wide, autocorrelating it four times to produce an approximation to a Gaussian function, truncating it to 21 pixels wide, and convolving it with a rectangular function 25 pixels wide. The window of Figure 6.11c was formed by taking a rectangular function 2 pixels wide, autocorrelating it eight times, and truncating the result to 45 pixels wide. The ZAA maps formed from the variously windowed speckle images are shown in the centres of Figures 6.11a-c. Comparing these ZAA maps with the ZAA map from ideal speckle images of the same object (Figure 6.4a) shows that the quasi-steadfast zero clusters have spread out in each case but that the migration of the clusters is quite different for the three windows. For the rectangular window (Figure 6.11a), the clusters migrated towards the u-axis. For the rectangle$\circledcirc$Gaussian window (Figure 6.11b) the clusters did not migrate much. For the Gaussian window (Figure 6.11c) the clusters migrated away from the u-axis. The zero-and-add images (right hand side of Figures 6.11a-c) reflect the direction and amount of migration of the clusters. The reconstruction errors for the three images are shown in Figure 6.11d. Both of the smoothly tapered windows caused considerably less degradation of the zero-and-add image than did the rectangular window. The rectangle$\circledcirc$Gaussian window is the best of the three, from both the visual similarity of $s_{ZA}(x)$ to the object and the comparatively low reconstruction error.

Figure 6.11 (facing page): The effect on zero-and-add of truncating speckle images with windows of rectangular (a), rectangle ⊙ Gaussian (b), and Gaussian (c) shape.

(a) - (c) (each) Left - Window function, of 45 pixels extent.
Centre - ZAA map from 300 statistically independent windowed speckle images of the object in Figure 6.2a.
Right - Zero-and-add image.

(d) Zero-and-add reconstruction errors in (a), (b), and (c).

The speckle images were generated as described in the caption to Figure 6.9 but with the window functions shown here instead of the rectangular windows of Figure 6.9.

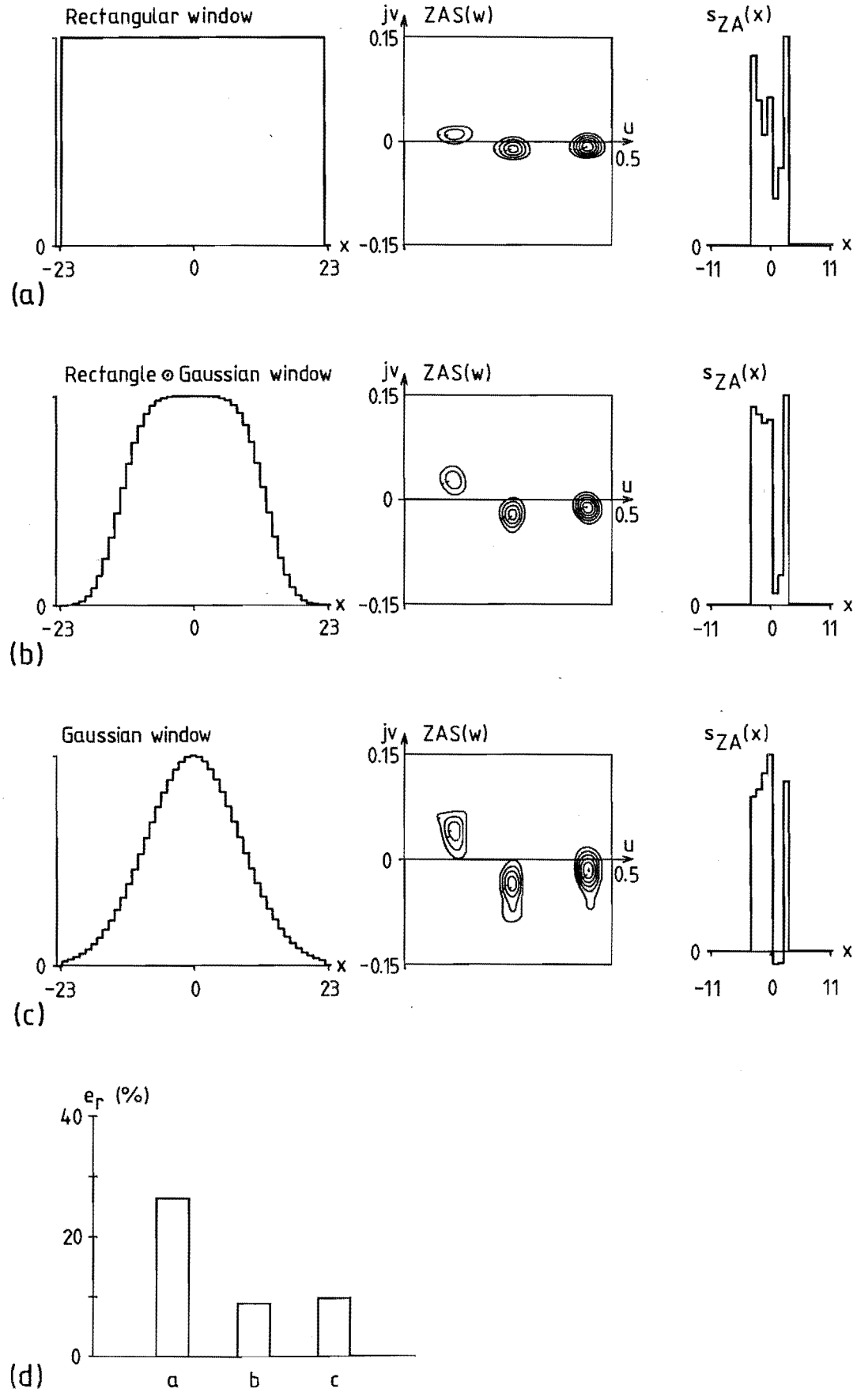

Rectangular window
jv
ZAS(w)
$s_{ZA}(x)$
0.15
0
-0.15
u
0.5
x
-23
0
23
-11
11
(a)
Rectangle ⊙ Gaussian window
(b)
Gaussian window
(c)
e_r (%)
40
20
0
a
b
c
(d)

Figure 6.11

6.6 CONCLUSIONS

The zero-and-add technique introduced in this chapter is a novel approach to speckle imaging. The principle of zero-and-add is founded on an isoplanatic description of speckle images. The practical algorithm described in Section 6.4 extends the principle to encompass speckle images with significant deviations from perfect isoplanatism.

The computational studies reported in Section 6.5 show that contaminating speckle images by adding pseudo-random noise, photon-limiting them, or truncating them with window functions causes the quasi-steadfast zeros in the ZAA map formed by zero-and-add to become diffuse and to migrate from their true positions in the complex plane. These effects become more pronounced as the contamination becomes more severe. The latter effect causes a gradual deterioration of the quality of the zero-and-add image with increasing contamination, and the former eventually causes a loss in quality when quasi-steadfast zeros become too diffuse to be correctly identified in the ZAA map. Additive noise with a probability density function that is uniform across the speckle image causes the quasi-steadfast zeros to migrate towards the real axis of the complex plane. This is the main cause of deterioration of the zero-and-add image when such additive contamination is applied. Photon limiting also causes the quasi-steadfast zeros to migrate towards the real axis. Its effects on zero-and-add are qualitatively similar to those of uniform additive noise.

Contaminating ideal speckle images by truncating them has a marked effect on the zero-and-add image. Truncation by rectangular windows degrades the zero-and-add image in a manner similar to that of uniform additive noise and photon limiting. In particular, the truncation makes the quasi-steadfast zeros in the ZAA map migrate towards the real axis of the complex plane. In the computational examples of Section 6.5.3, the amount of migration did not decrease uniformly as the width of the rectangular window increased. Thus the sensitivity of zero-and-add to truncation does not simply depend upon the size of the discontinuity at the boundary of the truncated speckle images or upon the proportion of the energy of the speckle images lost

in truncation. Window functions of other than rectangular shape can affect the quasi-steadfast zeros in the ZAA map quite differently from rectangular windows. In Section 6.5.3, a truncated Gaussian window caused the quasi-steadfast zeros to migrate away from the real axis instead of towards it, while a window formed by convolving a rectangle with a truncated Gaussian caused comparatively little migration and consequently yielded a zero-and-add image with little degradation. This result offers hope that window functions can be devised to minimise the impact of truncation on zero-and-add.

Perhaps the greatest potential advantage of zero-and-add over other speckle imaging methods is the possibility it offers of being able to generate faithful images of objects of arbitrary form with no need to compensate for the statistics of the seeing. This possibility will be realised if speckle images can be obtained in practice with sufficiently low contamination that (a) the quasi-steadfast zeros are steadfast enough to be recognisable immediately in the averaged zero map in the manner of the example presented in Figure 6.2, and (b) the migration of the quasi-steadfast zeros is negligible. Significant disadvantages of zero-and-add, compared to shift-and-add for example, are that it is computationally expensive and limited to objects of comparatively small extent. These disadvantages both stem from the numerical difficulty of factorising polynomials of high degree. The subroutines CPOLY and RPOLY used to locate the complex visibility zeros of speckle images (cf. ϕ6.2) tend to become numerically unstable for speckle images with extents greater than about 70 pixels. This restriction on the degree of polynomial that can be factorised means that speckle images may frequently have to be truncated in practice. Minimising the deleterious effects of truncation on zero-and-add is therefore particularly important.

The zero-and-add principle as introduced in this chapter applies only to data which are either one-dimensional or can be pre-processed into one-dimensional form. This is because only one-dimensional images have visibilities characterised by discrete zeros. One-dimensional speckle images may be formed with a telescope having a linear instead of a circular aperture [cf. Bates and Fright 1982]. Aime et al [1983] point out the practical advantages of such a telescope for speckle interferometry [cf. Sinton et al 198_ ϕ6]. The

one-dimensional telescope would also be well suited to zero-and-add imaging. Zero-and-add can be applied to two-dimensional speckle images obtained from a conventional telescope by computing one-dimensional projections of the speckle images and performing zero-and-add on the projections to form zero-and-add images of corresponding projections of the object. The zero-and-add images can usefully augment the images formed by other speckle processing techniques [Sinton et al 198_]. Faithful two-dimensional images can be reconstructed from the projections alone if the projections are sufficiently numerous and are themselves faithful [Lewitt and Bates 1978, Lewitt et al 1978], as has been demonstrated by the success of computed tomography [cf. Bates et al 1983] in medical imaging.

Part 2

IN MEDICINE

CHAPTER 7

TRANSDUCTION OF LUNG VENTILATION

7.1 INTRODUCTION

This chapter presents a brief review of the transduction and measurement of ventilation of the lungs, with emphasis on techniques which are suitable for the long-term monitoring of patients in hospitals.

In Section 7.2 the relationship between respiration, ventilation, and breathing is explained, the physiology of the lungs is briefly touched on, and the lung volume changes during breathing commonly referred to in the medical literature are defined. Section 7.3 describes transducers of gas volume and flow measured at the mouth and nose of the patient, i.e. by direct connection to the airway. In Section 7.4 the indirect techniques of measuring ventilation of the lungs from the movements of the torso are described. Inductance plethysmography and the calibration of systems which measure chest and abdomen motions are discussed in some detail. The requirements of long-term respiratory monitoring in a clinical setting are discussed in Section 7.5.

A medical glossary near the beginning of this thesis gives definitions for medical terms used in this chapter and in Chapters 8 and 9.

7.2 THE LUNGS AND LUNG VOLUMES

Living tissues require energy to drive the chemical processes by which they function. This energy is generated in the cells of the tissues by the combustion of carbon compounds with oxygen to form carbon dioxide, other chemical combustion products, and energy. In animals the carbon compounds which are oxidised in the cells for

energy come from the digestion of food. Oxygen is obtained from the atmosphere and the carbon dioxide formed is returned to the atmosphere. The exchange of oxygen and carbon dioxide between the tissues of the body and the atmosphere is termed respiration.

In humans the principal organs of respiration are the two lungs. These organs provide for the diffusion of gases between the atmosphere and a body fluid called blood. The blood is pumped by the heart to all parts of the body through the circulatory system [Guyton 1971 ϕϕ11-23]. It carries oxygen from the lungs to the body tissues and carbon dioxide from the body tissues to the lungs.

The lungs are spongy organs whose structure provides a large area through which gases can diffuse between the atmosphere and the blood. The lungs contain very many tiny chambers called alveoli through whose walls most of the gas exchange occurs. The alveoli are connected to the ends of a branching tree of air vessels whose stem (the trachea) connects to the atmosphere via the mouth and nose. For gas exchange between the blood and the air within the alveoli to continue, the alveolar air must be replenished by fresh air. Transfer of gases between the atmosphere and the alveoli is accomplished by a cyclical expansion and contraction of the lungs. This movement of the lungs and the associated body motions is called breathing. The transfer of fresh air into and out of the lungs is termed the ventilation of the lungs. The analogous supply of fresh blood to the lungs is called the perfusion of the lungs. The distinction between breathing, ventilation, and respiration is important and should be carefully remembered.

Breathing is driven by muscles in the chest wall and the abdomen. In vigorous breathing, e.g. during exercise, both inspiration (expansion of the lungs) and expiration (contraction of the lungs) are actively forced by contracting muscles. In quiet breathing, e.g. while resting or asleep, inspiration is driven by muscle contraction and expiration by the elastic recoil of lungs and thorax while the inspiratory muscles relax [Nunn 1977 pp.167-170].

Normal breathing comes from the involuntary contraction and relaxation of muscles which also may be controlled voluntarily. The

rhythm of involuntary breathing originates not from the muscles themselves, by comparison with cardiac rhythm which originates from the heart muscle [Guyton 1971 ϕ12], but from the brain [Nunn 1977 ϕ2]. The rhythm of breathing is affected by factors such as the state of consciousness and the degree of exertion of the individual, and may be interrupted by involuntary or voluntary acts such as coughing, sneezing, swallowing, and talking.

The amount of ventilation of the lungs depends on the frequency and depth of breathing. These are regulated by chemical control mechanisms within the body to maintain homeostasis (constant internal conditions in the presence of changing external conditions) against external changes such as altered composition of the inspired air and internal changes such as increased oxygen demand and carbon dioxide production from exercise [Nunn 1977 ϕ2].

Figure 7.1 shows a diagram of the lung volumes commonly quoted in studies of breathing. The tidal volume V_T is the amplitude of normal or "tidal" breathing. The vital capacity is the amplitude of a breath between the greatest possible inhalation and exhalation without forced or rapid effort. The functional residual capacity is the volume of the lungs (this is taken to include all of the airways) at the end of a normal expiration. The residual volume is the gas volume

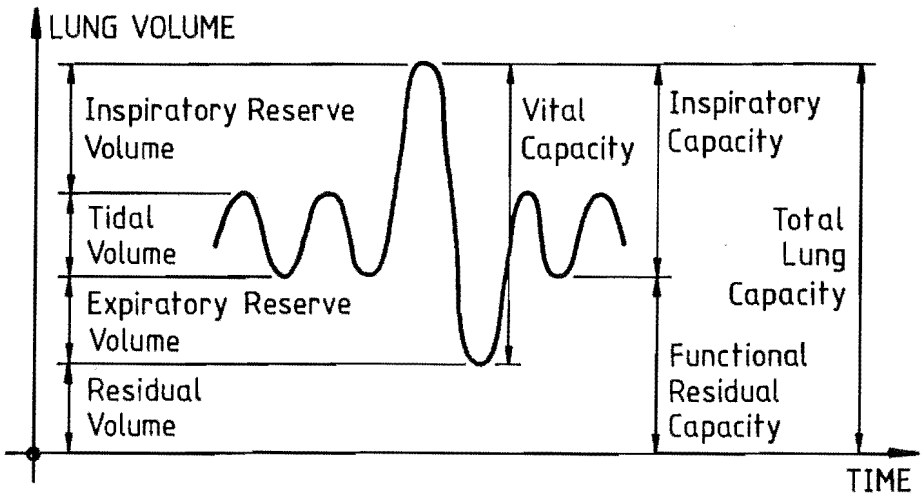


Figure 7.1: Names of static lung volumes and changes in lung volume (after Nunn [1977]).

which cannot be expelled from the lungs [Nunn 1977 pp.4-, Ruppel 1982 ϕ1]. Guyton [1971 p.329] points out that each of the "capacities" is the sum of two or more of the "volumes". The lung volumes of Figure 7.1 are called "static" because their definitions and measurement do not explicitly require particular histories of lung volume change, unlike dynamic tests of pulmonary mechanics and pulmonary function [Nunn 1977 ϕ4, Ruppel 1982 ϕ3].

The minute ventilation or minute volume $\dot{V}_E$ is the total volume of air expired, or inspired, in one minute [Tisi 1980 p.6, Ruppel 1982 ϕ2]. The subscript E stands for "expired", but the inspired and expired volumes are essentially equivalent. Since the minute ventilation is a volume over a given time it actually expresses a flow. The use of the dot-above in $\dot{V}_E$ indicates the time derivative of volume. For steady tidal breathing the minute ventilation is related to tidal volume and breathing rate f by

$$\dot{V}_E = V_T f \ . \qquad (7.1)$$

Typical values for tidal volume and breath period $\tau = 1/f$ in a normal resting adult are 500ml and 5 seconds respectively, giving a minute ventilation of 6 litres per minute [Guyton 1971 p.330, cf. Newsom Davis and Stagg 1975]. The minute ventilation expresses the rate at which inspired air is made available to the lungs for gas exchange with the blood. Inadequate ventilation impairs the respiration process [Nunn 1977 ϕ6].

A quantity which is particularly relevant when external apparatus is connected to the subject's airway (for whatever purpose) is the dead space V_D. This is the portion of the tidal volume which does not contribute to gas exchange with the blood. Dead space exists because air does not flow continuously through the lungs but instead flows in and out in a reciprocating manner along the same air passages. Since not all of the interior of the lungs and airways permits exchange of gases between the air within them and the blood, a certain part of the air drawn in with each breath comes to rest in locations (e.g. the large air passages, alveoli which are unperfused) where it does not contribute to gas exchange with the blood [Nunn 1977

ϕ7]. The volume of this part of the inspired air is the dead space. In tidal breathing only the part $(V_T - V_D)$ of each tidal volume is effective for gas exchange, and from this the effective part of the minute ventilation, called the alveolar ventilation $\dot{V}_A$, is given by

$$\dot{V}_A = (V_T - V_D)\, f \qquad (7.2)$$

(cf. 7.1). Connecting respiratory apparatus to the airway (e.g. by breathing through a piece of tubing) increases the dead space, necessitating an increase in minute ventilation if the alveolar ventilation is to be maintained constant [cf. Goode et al 1969].

7.3 MEASUREMENT OF VENTILATION BY CONNECTION TO THE AIRWAY

Lung ventilation is measured directly by connecting apparatus to the subject's airway. The connection can be made using a mouthpiece with a noseclip to block the nose, or a facemask covering both mouth and nose, or by enclosing the head or the entire body of the subject. Normally one desires quantitative information on lung volume changes (Figure 7.1) or gas flows during breathing. Transducers of various sorts exist for both gas volume (ϕ7.3.1) and gas flow (ϕ7.3.2). Volume changes can be deduced from flow measurements by mechanically, electrically, or numerically integrating the flow signal. Similarly, flow can be obtained by differentiating volume measurements. Nunn [1977 ϕ6] notes that flow transducers generally have smaller dead space and a broader frequency response than volume transducers but that spirometers (ϕ7.3.1) can measure gas volume more accurately than flow sensors can.

Access to the inspired and expired gases allows quantities other than gas volume and flow to be measured. For example, continuous analysis of the composition of the expired gas with a fast-responding gas analyser reveals cyclic variations of carbon dioxide concentration which are descriptive of breathing [Osborn 1977, Laxminarayan et al 1983].

7.3.1 MEASUREMENT OF RESPIRATORY VOLUME

The simplest transducer of respiratory volume is the spirometer, which is merely an extensible chamber into which the subject breathes (see Figure 7.2). Practical spirometers exist in a variety of sizes and types specialised for different applications. For general-purpose use such as determining tidal volumes or minute ventilation, water-seal spirometers consisting of a bell dipped into water have seen wide use in medical practice [Tisi 1980 ϕ4, Ruppel 1982 ϕ9]. Dry-seal spirometers with a wedge or piston design are now becoming popular. Spirometers provide output either as the position of a pen drawing onto paper on a drum (this arrangement is called a kymograph [Roth 1978]) or as electrical voltages.

Spirometers are able to directly measure changes in lung volume, but they cannot measure absolute lung volume because the lungs cannot be completely deflated. Thus, in Figure 7.1, spirometry can measure only the volumes or capacities which do not include the residual volume. More sophisticated means are required to measure lung volumes which do include the residual volume. By placing the

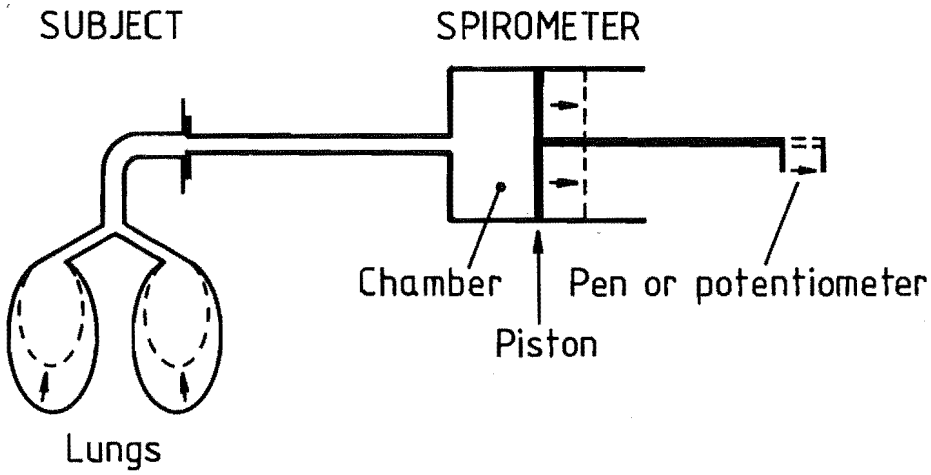


Figure 7.2: Principle of the spirometer: volume change in lungs plus volume change in spirometer equals zero. The subject is here shown exhaling into the spirometer. The spirometer produces an electrical or mechanical output dependent on the position of its piston, indicating changes in lung volume.

subject inside a body plethysmograph [Tisi 1980 p.94, Ruppel 1982 pp.141-142] and simultaneously measuring volume and pressure changes within the body plethysmograph and within the subject's airway during special breathing manoeuvres, Boyle's law may invoked to deduce the absolute lung volume [Ruppel 1982 pp.8-10]. Alternatively, analysis of gas composition after, say, introducing a known amount of a tracer gas to the gas being breathed allows the absolute lung volume to be determined [Nunn 1977 ϕ1, Ruppel 1982 ϕ1].

The simple spirometer arrangement of Figure 7.2 has the disadvantage of being a closed system in which the subject rebreathes the same gas. Unless precautions are taken, the carbon dioxide concentration rises in the rebreathed gas and causes pronounced changes in the subject's breathing. Eventually the subject would suffocate. Some means of removing carbon dioxide and replenishing oxygen in the air within the system must be provided before spirometry can be used for long-term or continuous measurement. This can be done by chemical means or by opening the system and providing a steady flow of fresh air through the spirometer [Spencer et al 1972, Nunn 1977 p.209].

The Wright respirometer is a mechanically integrating flow sensor. It consists of a small turbine connected through a gear train to a clock face on which integrated flow, i.e. volume change, is displayed [Wright 1955, Nunn 1977 p.210]. It responds to gas flow in one direction only, and thus may be breathed through continuously to measure ventilation. It has very little dead space and is sufficiently accurate for clinical applications [Nunn 1977 p.211].

Other types of gas meter which record gas volume are described by Hill [1973 pp.115-118], Nunn [1977 pp.209-210], and Sykes et al [1981 pp.186-187].

7.3.2 MEASUREMENT OF RESPIRATORY FLOW

Instruments which directly measure respiratory flow are collectively called pneumotachographs [Ruppel 1982 ϕ6]. Since respiratory flow is oscillatory and may vary rapidly with time,

instruments (i.e. pneumotachographs) designed primarily for measuring it are distinct from instruments (e.g. rotameters [Sykes et al 1981 pp.191-196]) whose main purpose is to measure steady gas flows.

Gas flow within a tube can be measured by partially obstructing the gas flow at some place within the tube and measuring the drop in pressure of the gas caused by its passage through the obstruction. This is analogous to measuring an electric current in a wire by measuring the voltage drop across an electrical resistance connected in series with the wire. The obstruction to gas flow is here called a flow resistor (see Figure 7.3). A popular flow transducer using this principle was introduced by Fleisch [1925]. The flow resistor in the Fleisch and similar pneumotachographs consists of a bundle of many parallel, narrow tubes [Hill 1973 p.109]. This arrangement is designed to promote laminar flow of the gas through the resistor so that, over a useful range of flow, the pressure drop is approximately linearly proportional to flow according to the Hagen-Poiseuille law [Nunn 1977 ϕ4]. The resistance of laminar-flow resistors depends on the viscosity of the gas, which in turn depends on factors such as its composition and temperature [Grenvik et al 1966]. To cover the complete range of flows encountered in medical practice a number of different-sized Fleisch transducers are used [Hill 1973 p.110].

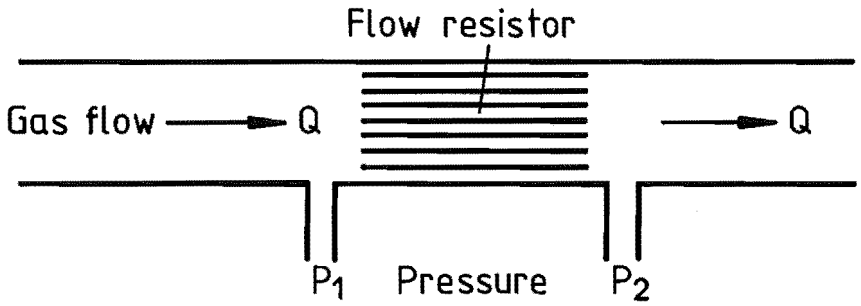


Figure 7.3: Flow measurement with a flow resistor. Passage of gas through the resistor is accompanied by a decrease in pressure of the gas. The pressure difference across the resistor is measured. For laminar flow through the resistor, the pressure difference is proportional to the gas flow, i.e. $\Delta P = P_1 - P_2 = kQ$.

Because laminar-flow resistors consist of many fine tubes or openings, they are sensitive to moisture and mucus in the gas stream [Osborn 1977] condensing or lodging on the resistance element and changing its resistance. The flow resistor may be heated to minimise condensation, though the heating itself affects the calibration of the pneumotachograph [Turney and Blumenfeld 1973]. Flow resistors employing large orifices and turbulent flow have been devised to give better performance in damp conditions. The relationship between pressure drop and flow is inherently nonlinear for turbulent flow [Elliott et al 1977]. To obtain an output approximately linearly proportional to flow, the instrument may be linearised using electronic circuitry [Elliott et al 1977] or by mechanical design in which the size of the orifice varies appropriately with flow [Osborn 1978].

A turbine or anenometer can be used to measure gas flow, as in the Wright respirometer [Wright 1955]. The speed of rotation of the turbine is ideally linearly related to gas flow. Gas leakage past the turbine and friction in the mechanism cause deviation from linearity, especially at low flow rates [Nunn 1977 p.211]. Friction can be minimised by using an optointerruptor instead of a mechanical coupling to sense the turbine's rotation [Cox et al 1974]. Electronic techniques may be used to further improve the linearity of the transducer [Crane and Stuttard 1976].

Hot thermistor pneumotachographs measure the rate of heat loss from a hot element which is cooled by the passing gas. Typically, the thermistor is maintained at constant temperature or at a constant temperature above the gas flowing past it [Lundsgaard et al 1979, Ruppel 1982 p.130]. The relationship between heat loss and gas flow depends on the temperature difference between the thermistor and the gas and on the composition of the gas [Lundsgaard et al 1979]. The thermistor may be physically small and have low thermal inertia, giving the transducer a wide frequency response (e.g. response times less than 2ms in the transducer described by Lundsgaard et al [1979]). For non-quantitative applications such as apnoea monitoring, the small size of hot thermistor transducers allows them to be used without a bulky airtight connection to the subject's airway, e.g. with a nasal catheter [Stark 1982].

Sykes et al [1981 p.200] and Ruppel [1982 p.130] describe a flow transducer in which vortices are created at an obstruction so that their rate of formation is proportional to gas flow. The vortices are detected ultrasonically. The transducer is said to be comparatively insensitive to gas composition, temperature, and humidity.

7.4 MEASUREMENT OF VENTILATION FROM BODY MOTIONS

The direct methods of measuring ventilation by connection to the subject's airway (ϕ7.3) are unsuitable in some situations. There are two reasons why this may be so. The first is the cumbersome nature of the connection to the airway. In long-term clinical monitoring, for example, the necessary equipment can be uncomfortable for the patient and may interfere unduly with the routines of patient care. The second reason is that connection of apparatus to the airway can change the patterns of breathing [Gilbert et al 1972, Weissman et al 1984]. This can be due partly to increased dead space [Sackner et al 1980c], partly to psychological effects (the subject breathes consciously) [Tabachnik et al 1981], and partly to stimulation of nerves of the face (in infants) [Dolfin et al 1983].

To avoid these difficulties, indirect techniques for measuring ventilation have been developed. These techniques measure the movements of the body associated with breathing. There is in general a compromise between the convenience or "non-invasiveness" of the techniques and their quantitative accuracy. The degree of compromise is chosen differently for different applications, ranging from noncontact systems for determining only the presence or absence of ventilation in apnoea monitoring [Franks et al 1976] through to the use of X-ray computed tomography to determine the geometry of the interior of the thorax [Hendenstierna et al 1985].

7.4.1 REVIEW OF TECHNIQUES

Transducers of ventilation optimised for convenience are used in long-term monitoring applications where quantitative measurements

of ventilation are not required. An example of this is apnoea monitoring of infants, where only the presence or absence of ventilation need be determined. If apnoea, i.e. cessation of ventilation, is detected for a sufficiently long period then an alarm is sounded or some other action is taken automatically. Systems for respiratory monitoring of infants, that do not require physical contact with the infants, have been reviewed by Franks et al [1976]: the transducers include an air mattress on which the infant lies and by breathing causes pressure changes or air movements in the mattress, electrode pairs placed near the infant whose movements affect the interelectrode capacitance, permanent magnets attached to the infant, and radar. Radar systems have been calibrated quantitatively for tidal volume with adult subjects in a fixed position [Ambrosino and Bramanti 1983]. Air flow through the nose may be detected acoustically [Stark 1982]. A small cup applied to the anterior abdomen wall experiences changes of interior pressure related to breathing motions [Wright 1977].

Body movement sensors which have been calibrated to give a usefully accurate quantitative estimate of lung volume changes can be broadly classified as transducers of body diameter, circumference, cross-sectional area, or three-dimensional geometry. Transducers of each type are described in the following paragraphs.

Diameter measurements are normally made of the anteroposterior (front-to-back) diameters of the torso. Konno and Mead [1967] initially measured these diameters mechanically using strings attached to the subject. Subsequently they and others have employed magnetometers [Mead et al 1967], in which one of a pair of coils affixed to the front and back of the torso is energised with alternating current and the induced voltage in the other coil indicates the separation between the coils. The radar system of Ambrosino and Bramanti [1983] measures changes in anteroposterior diameters. Robertson et al [1980] report improved accuracy in determining lung volume changes by measuring lateral (side-to-side) diameters as well as anteroposterior diameters. The additional measurements give information on changes in cross-sectional shape of the torso.

Measurements of circumference are made using strain gauges attached to elastic material stretched around the torso [Faithfull et al 1979]. A commonly used transducer of this type is the mercury-in-rubber strain gauge, in which the electrical resistance of mercury contained within rubber tubing changes as the tubing is stretched longitudinally [Geddes and Baker 1975 ϕ2.10, cf. Sumner 1978b].

The respiratory inductance plethysmograph is a comparatively recent development [Milledge and Stott 1977, Cohn et al 1978]. The self-inductance of an elastic coil placed around the torso gives a measure of the cross-sectional area of the torso. The transduction of cross-sectional area allows more accurate estimates of lung volume changes than the measurement of anteroposterior diameters does, in the presence of changes of cross-sectional shape of the torso such as occur with changes of posture [Sackner et al 1980b]. Inductance plethysmography is described in more detail in Section 7.4.2, and an inductance plethysmograph built by myself is described in Chapter 8.

In electrical impedance plethysmography the electrical impedance between electrodes attached to the body surface is measured, giving information about changes in volume, geometry, or composition of those parts of the body through which the current flows [Nyboer 1970, Geddes and Baker 1975 ϕ10, cf. Seagar 1983a] (see also Chapter 9). Hill [1973 pp.118-125] and Geddes and Baker [1975 ϕ10.13] describe impedance plethysmographic measurement of breathing. The origin of the impedance changes during breathing is not completely understood [Kira et al 1971, Geddes and Baker 1975 ϕ10.13], because the spatial distribution of current flow within the subject is poorly known. Electrode placement and arrangement have an important effect on the linearity [Geddes and Baker 1975 ϕ10.13] and sensitivity to motion artefacts [Meijer et al 1982, Itoh et al 1982a, Sahakian et al 1985] of the impedance plethysmograph. The relationship between impedance changes and lung volume changes can be significantly nonlinear [Geddes and Baker 1975 ϕ10.13, Itoh et al 1982a]. Cardiac activity causes artefactual impedance changes [Wilson et al 1982b]. Independent measurements of motion of different parts of the body may be made using two sets of electrodes operating at different frequencies [Hill 1973 p.123]. Electrodes may be shared with ECG

monitoring equipment. Impedance plethysmography has been used as the primary transducer of breathing in at least two computerised respiration analysis systems described recently [Wilson and Franks 1982, Itoh et al 1982b].

7.4.2 INDUCTANCE PLETHYSMOGRAPHY

Inductance plethysmography is a non-invasive technique for measuring the cross-sectional area of a part of the body. The inductance plethysmograph transducer is a coil of conductive wire attached to an elastic band. When the band is placed around a part of the body (e.g. the torso), the wire is held close to the body surface. The cross-sectional area of by the coil is then approximately equal to the cross-sectional area of the body where the coil is. The electrical inductance of the coil is measured. The inductance of the coil is strongly dependent on its cross-sectional area but, for the rounded shapes typical of cross-sections of the body, only weakly dependent on its shape (see ϕ8.4.3.2). Thus, measuring the inductance of the coil provides an accurate estimate of the area of the cross-section enclosed by the coil, even though the exact shape of the cross-section is unknown.

Inductance plethysmography as a transducer of ventilation was first described by Milledge and Stott [1977] and by Cohn et al [1978]. Subsequent general descriptions of respiratory inductance plethysmography (which is often abbreviated to RIP in the literature) are given by Watson [1979], Sackner et al [1980b], and Cohn et al [1982]. The respiratory inductance plethysmograph of Milledge and Stott [1977] used a single tall coil placed around the torso, extending from the upper chest to the pubis. The coil formed part of the resonant circuit of an oscillator, and the coil inductance was measured by measuring the frequency of oscillation. The coil was affected by breathing motions of both the chest and the abdomen. Used as a transducer of respiratory volume, this system was reported to be linear against a spirometer to within 2% over most of the vital capacity range, but was sensitive to motion artefacts. More recent inductance plethysmographs have used two separate coils, worn over the chest and over the abdomen [Watson 1979, Sackner et al 1980b], and are

sufficiently insensitive to motion artefacts to be useable during exercise [e.g. Sackner et al 1980c]. The two coils provide separate measurements of the cross-sectional areas of the rib cage and abdomen. The signals from the coils are multiplied by individual weighting factors and added to provide an estimate of respiratory volume (ϕ7.4.3). Respiratory inductance plethysmographs using the two-coil arrangement are commercially manufactured in the USA under the trade name "Respitrace".

Studies of the accuracy of respiratory inductance plethysmography as a transducer of ventilation have been made on normal adults [Hill et al 1982, Chadha et al 1982, Zimmerman et al 1983, Owen et al 1985], adults with pulmonary disease [Tobin et al 1983a, Gonzalez et al 1984], children [Tabachnik et al 1981], infants [Duffty et al 1981, Dolfin et al 1982, Warren and Alderson 1985], and sheep [Watson et al 1981, Abraham et al 1981]. The average magnitude of error between inductance plethysmography and spirometry in measuring the tidal volumes of breaths of adult subjects in several postures has been variously reported as roughly 5% [Gonzalez et al 1984], 10% [Zimmerman et al 1983], or up to 15% [Owen et al 1985].

Inductance plethysmography has been used to study respiratory patterns during rest and exercise in healthy adult subjects [Sackner et al 1980a, Sackner et al 1980b, Tobin et al 1983b] and sufferers of pulmonary disease [Delgado et al 1982, Tobin et al 1983c, Sackner et al 1984a, Sackner et al 1984b], breathing by asthmatics [Chadha et al 1981, Warren and Alderson 1984], respiratory changes induced by direct transducers of ventilation (ϕ7.3) [Sackner et al 1980c, Tabachnik et al 1981, Dolfin et al 1983], respiratory patterns after surgery [Jordan et al 1980, Catley et al 1982a, Catley et al 1982b], respiratory effects of anaesthetic drugs [Catling et al 1980, Royston et al 1981], the smoking of cigarettes [Tobin and Sackner 1982], and pulmonary mechanics in sheep [Abraham et al 1981].

7.4.3 THE CHEST/ABDOMEN MODEL OF BREATHING

Changes of lung volume during breathing occur in association with changes in volume of the torso as a whole. If the volume changes

of compressible regions of the interior of the torso other than the lungs are neglected, then changes in lung volume are equal to changes in volume of the entire torso, which may be measured from the movements of the body surface. Measurements of body movements at a single place on the torso are inadequate to accurately determine lung volume changes because the torso changes shape as well as volume during breathing. However, Konno and Mead [1967] observed that the volume changes of the torso are well described by measuring the separate movements of the chest (or rib cage) and the abdomen. They introduced a two-compartment model of breathing in which the volume change ΔV of the lungs is equal to the sum of volume changes ΔV_C, ΔV_A of chest and abdomen compartments, respectively, of the torso, i.e.

$$\Delta V = \Delta V_C + \Delta V_A \ . \tag{7.3}$$

This can be related to a physical picture of the torso (see Figure 7.4) by rewriting (7.3) as

$$V = V_C + V_A - c_1 \tag{7.4}$$

using the absolute volumes V, V_C, V_A of lungs, chest, and abdomen.

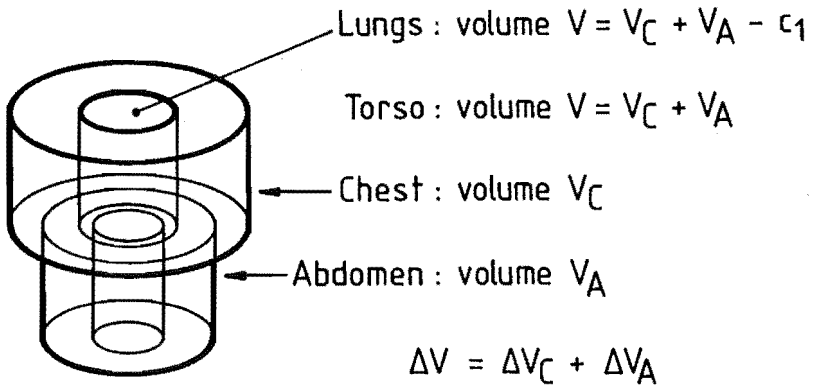


Figure 7.4: A physical picture of the two-compartment model of lung volume changes. Lung volume change ΔV is equal to the sum of volume changes ΔV_C and ΔV_A of chest and abdomen compartments. ΔV_C and ΔV_A may be inferred from measurements of chest and abdomen motions.

The constant c_1 represents the volume of the contents of the torso other than the lungs.

The volume changes ΔV_C, ΔV_A of the chest and abdomen compartments are characterised by measuring the movements of the chest and abdomen, using transducers of diameter, circumference, or cross-sectional area (ϕ7.4.1). The transducers are applied to the chest and to the abdomen and produce output signals, x and y say, which are usually electrical voltages that may be displayed on an oscilloscope, plotted with a chart recorder, or digitised by a computer. Provided the relationship between the volume of the compartment and the output of the transducer applied to it is smoothly curved (cf. Mead et al [1967], ϕ8.4.3.1 of this thesis), the relationship is approximately linear over the small relative changes of x, y, and V that occur during breathing. Thus one may write

$$V \approx ax + by + c \qquad (7.5)$$

where c is a constant distinct from c_1 in (7.4). From (7.5), changes ΔV in lung volume are given by

$$\Delta V \approx a\Delta x + b\Delta y \ . \qquad (7.6)$$

In respiratory monitoring one is usually concerned with lung volume changes, such as the tidal volume and vital capacity (Figure 7.1), so that (7.6) applies. The constants a and b are sometimes called volume-motion coefficients [Zimmerman et al 1983, Gribbin 1983]. In other parts of this thesis they are also called weighting factors or calibration coefficients. A calibration procedure must be invoked to determine the values of a and b before lung volume changes can be quantitatively estimated using (7.6) from measurements of chest and abdomen movements.

7.4.4 CALIBRATION OF CHEST/ABDOMEN MOTION SENSORS

Calibration of chest/abdomen motion sensors to allow quantitative measurement of lung volume changes using (7.6) involves determining the values of the volume-motion coefficients a and b.

Their values can in principle be calculated from knowledge of the geometry of the torso, but in practice they are almost always obtained from simultaneous measurements of Δx, Δy, and ΔV with the body motion transducers and a direct respiratory volume transducer such as a spirometer or integrated pneumotachograph (§7.3). The result of the measurements is a data set $\{\Delta x_i, \Delta y_i, \Delta V_i\}$ from which a and b in (7.6) may be computed.

A number of calibration procedures for chest/abdomen motion sensing systems have been described [e.g. Konno and Mead 1967, Stagg et al 1978, Chadha et al 1982, Gribbin 1983, Warren and Alderson 1985]. These procedures differ in the practical details of how the data set is obtained and in the mathematical details of how values for a and b are computed from it. The practical details of measurement of the data set, which determine the suitability of the values within it for solving for a and b, crucially affect the accuracy of the calibration, i.e. the error in subsequent estimation of lung volume changes from the body movement measurements. A least-squared-error solution for a and b is used below to illustrate how this is so.

Let $\Delta\hat{V}_i$ be an estimate of the true lung volume change ΔV_i corresponding to the motion transducer signal increments Δx_i and Δy_i, i.e.

$$\Delta\hat{V}_i = a\Delta x_i + b\Delta y_i \qquad (7.7)$$

(cf. 7.6). The error in the i^{th} estimate is

$$e_i = \Delta\hat{V}_i - \Delta V_i \ . \qquad (7.8)$$

The sum of the squared errors over the whole data set can be used as an error metric E describing the correctness of the values a and b,

$$E = \sum_i e_i^{\,2} \ . \qquad (7.9)$$

Choosing a and b so that E is minimised yields the least-squares solution (a,b), which from (7.7), (7.8), and (7.9) is given by

$$a = \frac{\sum \Delta y_i^2 \sum \Delta x_i \Delta V_i - \sum \Delta x_i \Delta y_i \sum \Delta y_i \Delta V_i}{\sum \Delta x_i^2 \sum \Delta y_i^2 - (\sum \Delta x_i \Delta y_i)^2} , \qquad (7.10a)$$

$$b = \frac{\sum \Delta x_i^2 \sum \Delta y_i \Delta V_i - \sum \Delta x_i \Delta y_i \sum \Delta x_i \Delta V_i}{\sum \Delta x_i^2 \sum \Delta y_i^2 - (\sum \Delta x_i \Delta y_i)^2} . \qquad (7.10b)$$

(7.10a) and (7.10b) each have the denominator

$$d = \sum \Delta x_i^2 \sum \Delta y_i^2 - (\sum \Delta x_i \Delta y_i)^2 . \qquad (7.11)$$

Suppose that, during the measurement of the data set $\{\Delta x_i, \Delta y_i, \Delta V_i\}$, the chest and abdomen compartments make the same relative contributions to all volume changes ΔV_i. An X/Y plot of y_i against x_i (where x_i and y_i are the transducer signal values from which Δx_i and Δy_i are obtained), with arbitrary zeros, is then a straight line (see Figure 7.5a)

$$\Delta y_i = k \Delta x_i \quad \forall\, i \qquad (7.12)$$

with the constant k being the slope of the line. The ratio $\Delta y_i / \Delta x_i$, hereafter called the chest/abdomen ratio, is constant. For these data the numerators and the denominator of (7.10) are all zero. There is no unique solution for a and b. The torso is actually behaving as one compartment rather than two, because the chest and abdomen signals are linearly dependent on each other. Only a single signal, which may be any linear combination of x and y, is required to characterise the lung volume changes ΔV. Given the two signals x and y there are infinitely many solutions for a and b which minimise the error metric E. The solutions are all equally satisfactory provided the chest/abdomen ratio remains exactly constant.

During natural, spontaneous breathing the chest/abdomen ratio varies somewhat [Stagg et al 1978]. A plot of y_i against x_i shows deviations from linearity (see Figure 7.5b). The variations in chest/abdomen ratio make a unique solution for a and b from (7.10) possible, because the denominator d (7.11) is nonzero (cf. Stagg et al [1978] who use a different mathematical procedure to obtain their

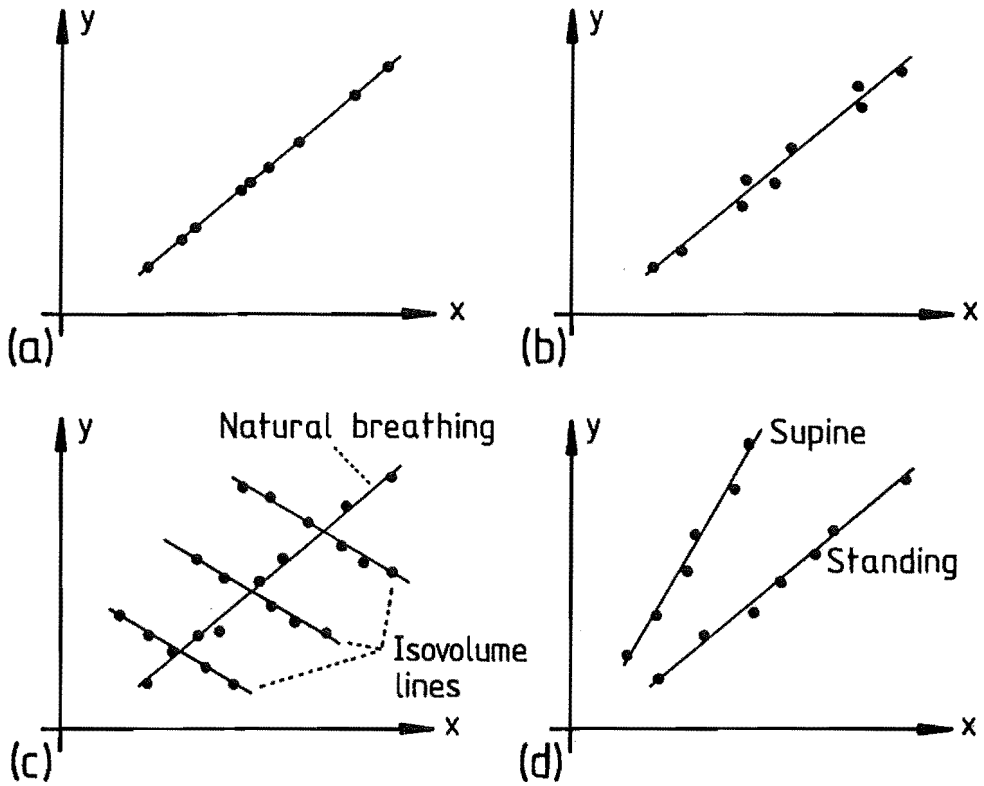


Figure 7.5: X/Y plots of simultaneous measurements of chest motion signal x and abdomen motion signal y. The fatness of the region covered by the measured data points shows how well-conditioned the solution for the volume-motion coefficients is (see text).

(a) Hypothetical situation. All data points lie on a straight line, $\Delta y/\Delta x$ = constant. There is no unique solution for the volume-motion coefficients.

(b) Natural breathing. The data points lie in a narrow region close to a straight line. The solution for the volume-motion coefficients is unique but ill-conditioned.

(c) Isovolume manoeuvres. Data points are spread over a broad region. The solution for the volume-motion coefficients is well-conditioned.

(d) Two-posture calibration. Changes in posture alter breathing so that the measured data collectively cover a broad region, giving a well-conditioned solution for the volume-motion coefficients.

solution). The value of d affects the mathematical conditioning of the solution, i.e. how little the values of a and b change in response to perturbations of the data Δx_i, Δy_i. The larger d is, the better conditioned is the solution, i.e. a and b change less when the data are perturbed. With a better-conditioned solution one can be more confident that the values of the volume-motion coefficients obtained from potentially noisy data are close to their true values for the geometry of the torso. The denominator d is increased by increasing the variation of the chest/abdomen ratio within the data set. On an X,Y plot (Figure 7.5) this increases the fatness of the smallest convex curve circumscribing all of the data points. The breathing of subjects with respiratory abnormalities such as obstructive pulmonary disease can exhibit greater than normal variability of chest/abdomen ratio, associated with the occurrence of obstructed and paradoxical breathing motions. The increased fatness of X/Y plots from diseased subjects compared to X/Y plots from normal subjects [Sackner et al 1984b] is evidence of this.

The variability of the chest/abdomen ratio within the data set can be increased by ensuring that the breathing pattern of the subject changes during the measurement of the data set. This has been done in several ways in practice. The earliest and still a popular technique is the isovolume manoeuvre [Konno and Mead 1967]. The subject's airway is occluded, by closing the glottis or by blocking a breathing-tube connecting the subject's airway to (say) a spirometer, and the subject transfers air to and fro between the chest and abdomen compartments. Signals x and y are measured for one or several constant values of V. Measurements may also be made of x, y, and V during normal breathing. The measurements (see Figure 7.5c) collectively cover a broad region on the X,Y plot, ensuring that d is large and that the solution for a and b from (7.10) is well-conditioned. For isovolume manoeuvres at several known increments of lung volume, the solution for a and b may be obtained graphically from plots like Figure 7.5c [Konno and Mead 1967, Zimmerman et al 1983].

The biggest disadvantage of the isovolume manoeuvre is that it requires the cooperation of a trained and skilful subject. The air within the lungs must not be compressed significantly during the manoeuvre, otherwise the isovolume condition no longer holds. Various

means of inducing involuntary variations in the chest/abdomen ratio have been devised to allow calibration without special respiratory manoeuvres such as the isovolume manoeuvre. A tight band constraining the chest or abdomen motions has been used in measurements on sheep [Abraham et al 1981]. For clinical measurements on human patients, the changes in chest/abdomen ratio occurring spontaneously with changes in posture have been used and studied [Watson et al 1981, Tabachnik et al 1981, Chadha et al 1982, Zimmerman et al 1983, Tobin et al 1983a, Gonzalez et al 1984]. For calibration on infants, changes in chest/abdomen ratio in a single posture are obtained from changes in the state of consciousness of the subject [Duffty et al 1981, Dolfin et al 1982] or by selecting particular breaths for inclusion into the data set [Warren and Alderson 1985]. In each of these cases a data set $\{\Delta x_i, \Delta y_i, \Delta V_i\}$ including measurements with two (or more) significantly different chest/abdomen ratios (see Figure 7.5d) is sufficient to give a large denominator and hence a well-conditioned solution for a and b from (7.10).

The uncertain values obtained for the volume-motion coefficients when calibration is attempted with a data set yielding a degenerate or ill-conditioned solution (as in Figures 7.5a and 7.5b) do not lead to inaccurate estimates of lung volume changes in subsequent measurements using (7.6), unless the chest/abdomen ratio in the subsequent measurements becomes significantly different from what it was during the calibration measurements. Thus, any of the possible solutions from a data set obeying (7.12) give accurate results provided the subsequent breathing also obeys (7.12) with the same value of k. Calibrations made with a particular breathing pattern (e.g. with the subject in a particular posture) give accurate results provided the breathing pattern does not change [Duffty et al 1981, Warren and Alderson 1985]. When the breathing pattern does change, so that subsequent measurements are made over a range of chest/abdomen ratios, the accuracy of the calibration deteriorates significantly [Owen et al 1985]. This may be partly due to the ill-conditioned solution giving inaccurate values for the volume-motion coefficients, and partly due to dependence of the true values of the volume-motion coefficients on posture [Zimmerman et al 1983]. The latter effect arises because the two-compartment model described in Section 7.4.3 does not perfectly represent lung volume changes. Improved accuracy

of calibration can be obtained by including a wide range of chest/abdomen ratios in the calibration measurements [Gonzalez et al 1984]. Gonzalez et al [1984] claim that this yields "compromise" values for the volume-motion coefficients when the volume-motion coefficients do vary with posture.

7.5 DISCUSSION

I have been concerned with long-term monitoring of the lung ventilation of patients in hospital after surgery. The aim has been to investigate the respiratory effects of surgery and anaesthesia (see Chapter 8). In this application the transducer of ventilation employed in the monitoring apparatus should be able to give quantitative measurements of ventilation with adequate accuracy, but must not cause undue discomfort or inconvenience to the patient and the clinical staff. Direct transducers of gas flow (ϕ7.3) are in general unsuitable because they are too cumbersome and induce respiratory changes by their presence. Of the body-movement transducers (ϕ7.4), inductance plethysmography and impedance plethysmography are suitable for clinical respiratory monitoring. They do not excessively encumber or constrain a patient confined to a hospital bed. Both techniques have been used successfully in clinical settings [Catley et al 1982b, Itoh et al 1982b].

Body-movement transducers which separately measure the movements of chest and abdomen have an advantage over ventilatory transducers (e.g. spirometers) which measure ventilation directly (ϕ7.3), because the separate motion measurements allow clinically significant abnormalities of breathing such as paradoxical motions (see ϕ8.6.2.7) to be detected easily. With a spirometer these motions would difficult to distinguish from shallow but otherwise normal breathing. The two-coil respiratory inductance plethysmograph (ϕ7.4.2) can conveniently provide chest and abdomen measurements.

Accurate calibration of chest/abdomen body movement sensors for respiratory volume is likely to be more difficult in clinical settings than in the laboratory setting from which most published descriptions of calibration procedures (ϕ7.4.4) have come. This is

because hospital patients may not be able to perform special respiratory manoeuvres or have their postures deliberately changed to obtain variations in chest/abdomen ratio. The reference transducer of ventilation used to calibrate the body-motion transducer should not be excessively cumbersome, since the clinical setting demands that the respiratory monitoring equipment be mobile and small enough to be used in the sometimes cramped conditions of hospital wards. This suggests that a pneumotachograph would be preferable to a spirometer (ϕ7.3) as the reference transducer because of its small physical size.

The respiratory monitoring system described in Chapter 8 uses inductance plethysmography to non-invasively measure lung ventilation according to the two-compartment model of breathing described in Section 7.4.3. The points made in the previous paragraph and the contents of Section 7.4.4 are pertinent to the use of the system for quantitative measurement of ventilation in clinical practice.

CHAPTER 8

A COMPUTERISED SYSTEM FOR LONG-TERM RESPIRATORY MONITORING AND ANALYSIS

8.1 INTRODUCTION

This chapter describes a computerised system for long-term respiratory monitoring and analysis constructed by myself with the assistance of the University Division of Anaesthesia of the Christchurch Clinical School of Medicine. The system employs a non-invasive, indirect transducer of respiratory activity which allows routine use in the hospital wards without greatly interfering with normal procedures of patient care. A microcomputer automates the functioning of the system, monitoring and recording the respiratory movements breath-by-breath and having the potential for automated on-line data analysis.

The respiratory monitoring system was developed to allow the Anaesthetics Department to study the respiratory effects of opioid drugs administered to patients for pain relief after surgery. The topic of pain relief and the respiratory effects of opioid pain-killers is introduced in Section 8.2. Section 8.3 gives a brief overview of the respiratory monitoring system as it is currently being used at the Christchurch Hospital. The respiratory transducer, which uses the recently developed technique of inductance plethysmography (ϕ7.4.2), is described in Section 8.4. The microcomputer itself is described in Section 8.5. The data collection and data processing algorithms implemented by computer software for the respiratory monitoring system are described in Section 8.6. The data processing algorithms have been deliberately designed to be simple so that they can eventually be executed by the system's microcomputer while the measurements are being made (this real-time data processing has not yet been implemented). Initial clinical results obtained with the system are presented in Section 8.7. These indicate some of the changes in respiratory patterns that may occur as a result of opioid analgesia, and suggest that there is a significant difference between

the effects of general and regional analgesia. The performance and capabilities of the respiratory monitoring system, and potential applications outside postoperative monitoring, are assessed in Section 8.8. Definitions of medical terms used in this chapter may be found in the medical glossary near the beginning of this thesis.

8.2 POSTOPERATIVE PAIN RELIEF

8.2.1 PAIN RELIEF IN SURGERY

The usefulness of surgery as a treatment for disease, injury, or deformity is very much dependent on the adequate control, both during and after surgery, of the intense pain generated by the surgical wound [Wallace and Norris 1975]. Pain relief for the patient is one of the responsibilities of the anaesthetist, who works alongside the surgeon as part of the surgical team [Dripps et al 1982 ϕ1]. The principal tasks of the anaesthetist are to safeguard the welfare of the patient during surgery and to provide optimal operative conditions for the surgeon. The discipline of anaesthesia has a number of facets, including pain relief (for the patient's sake) and controlling muscle tension (which helps the surgeon). Pain relief has the special name analgesia.

The conventional and most widely used means of pain relief in Western medicine is the administration of drugs [Goodman Gilman et al 1980], which are chemical compounds used to alter the chemistry of the body's functioning in desired ways. Pain-killing or analgesic drugs interfere with the sensation of pain in response to normally painful stimuli. The opioid analgesics [Jaffe and Martin 1980], an important class of analgesic drugs, are widely used in controlling postoperative pain. The name "opioid" refers to drugs with morphine-like properties. Morphine was originally extracted from the juice of the opium poppy and is the archetypal opioid drug. These drugs act on the central nervous system [Thorpe 1984] to produce a number of effects including analgesia, narcosis (or stupor, hence they are also called narcotic drugs), and respiratory depression (ϕ8.2.2). A property of morphine-like drugs is that they confer analgesia without causing loss

of consciousness [Jaffe and Martin 1980], unlike some other anaesthetic drugs used during surgery [Dripps et al 1982] which render the patient unconscious.

There are various ways to administer analgesic drugs for postoperative pain relief. Opioid drugs may be injected into muscle tissue (intramuscular injection), injected close to the spinal cord (e.g. epidural and intrathecal injection) [Goodison 1983], infused into the blood stream [Church 1979], or taken orally [Derbyshire et al 1985], to produce analgesia. Alternatively, regional nerve blocks (e.g. the intercostal block) [Dripps et al 1982 ϕ21] may be used to produce analgesia over limited regions of the body.

8.2.2 RESPIRATORY EFFECTS OF OPIOID ANALGESIA

A side-effect of opioid analgesics, and many other drugs used in anaesthetic practice, is respiratory depression [Jordan 1982]. The drugs interfere with the rhythm of breathing and the chemical regulation of ventilation (ϕ7.2). The effects on the rhythm of breathing include decreased breathing rate and minute ventilation (ϕ7.2) [e.g. Rawal and Wattwil 1984, Clergue et al 1984] and the occurrence of short-term irregularities such as paradoxical chest and abdomen movement with obstruction to gas exchange [Catley et al 1982a, Catley et al 1982b, Jordan 1982]. Modification of chemical regulation leads to increased CO_2 concentration in the expired gases and decreased response of the respiratory centres [Nunn 1977 ϕ2] to changes in the oxygen and carbon dioxide content of the inspired gases [e.g. Weil et al 1975, Jordan 1982]. Respiratory depression is dangerous if it causes a significant reduction in the amount of oxygen carried by the blood from the lungs to the tissues of the body. When coupled with other factors such as pre-existing respiratory disease which impair the efficiency of oxygenation of the blood, the respiratory depression induced by opioid analgesia is potentially life-threatening [Dripps et al 1982 pp.442-].

To investigate the respiratory effects of analgesic drugs, it is useful to continuously monitor respiratory performance postoperatively because this allows sporadic effects of respiratory

depression to be reliably detected. It also allows the time history of respiratory performance to be related to, say, the times of administration of doses of the analgesic. Traditional measurements of respiratory performance such as assessment of ventilatory response to carbon dioxide or analysis of blood samples [Jordan 1982] can be performed only intermittently and are unsuited to routine use in the hospital wards because they are inconvenient and uncomfortable for the patient. Measurement of respiratory gas volume and flow by direct connection to the patient's airway [Ruppel 1982 ϕ9] also can be uncomfortable for the patient. Convenient, non-invasive, continuous means of monitoring have therefore been sought.

Blood gas content, specifically blood oxygen, can be continuously monitored by the technique of ear oxymetry [Merrick and Hayes 1976], in which light is shone through the lobe of the ear and the oxygen content of the blood is inferred from the light absorption from the oxygen carrying pigment haemoglobin at different wavelengths. Ventilation of the lungs is conveniently monitored continuously by measuring the body motions associated with breathing [Gribbin 1983]. Inductance plethysmography (ϕ7.4.2) is an effective way to do this. A recent study of respiratory effects of analgesic drugs [Catley et al 1982a, Catley et al 1982b, Jordan 1982] in which continuous postoperative monitoring of respiration was performed by ear oxymetry and inductance plethysmography revealed sporadic episodes of reduced blood oxygen associated with obstructed breathing in patients given opioid analgesia for postoperative pain relief.

The respiratory effects of different types of analgesia, such as different drugs, different dose size and timing, and different routes of administration, are not necessarily the same. Those analgesic techniques which cause the least respiratory depression are, in respect of their respiratory effects, the safest for the patient. Comparisons of the respiratory effects of different types of analgesia [e.g. Catling et al 1980, Shulman et al 1984, Choi et al 1985] can be valuable guides toward improving the quality of patient care in hospitals, since they allow respiratory effects to be included in the assessment of the safest and most suitable analgesia for particular situations. The respiratory monitoring system described in the following sections of this chapter was designed and constructed to be

a research tool for use in investigations of the respiratory effects of analgesic drugs.

8.3 THE RESPIRATORY MONITORING SYSTEM - OVERVIEW

The respiratory monitoring system (see Figure 8.1) currently in use at the Christchurch Hospital consists of an inductance plethysmograph (ϕ8.4) to measure the breathing motions of the patient, a microcomputer (ϕ8.5) to automate data collection and data analysis (ϕ8.6) by the system, and a visual display unit (VDU) for communication between the microcomputer and the operator of the system. The system lives on a trolley (Figure 8.2a) which makes it mobile within the hospital.

In clinical use, the system is brought on its trolley to the bedside of the patient in the hospital ward. The belts of the inductance plethysmograph are fitted to the patient (Figure 8.2b) and are connected to the plethysmograph by a cable when the patient is in bed. The microcomputer samples the electrical outputs of the plethysmograph at regular intervals (usually of one second; see

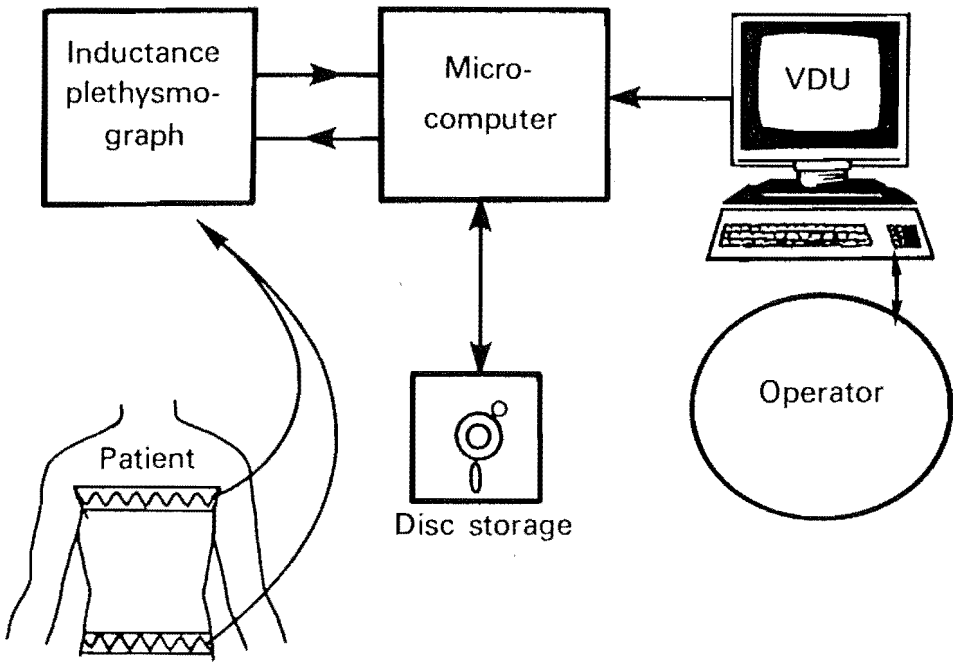


Figure 8.1: Block diagram of the respiratory monitoring system.

(a)

(b)

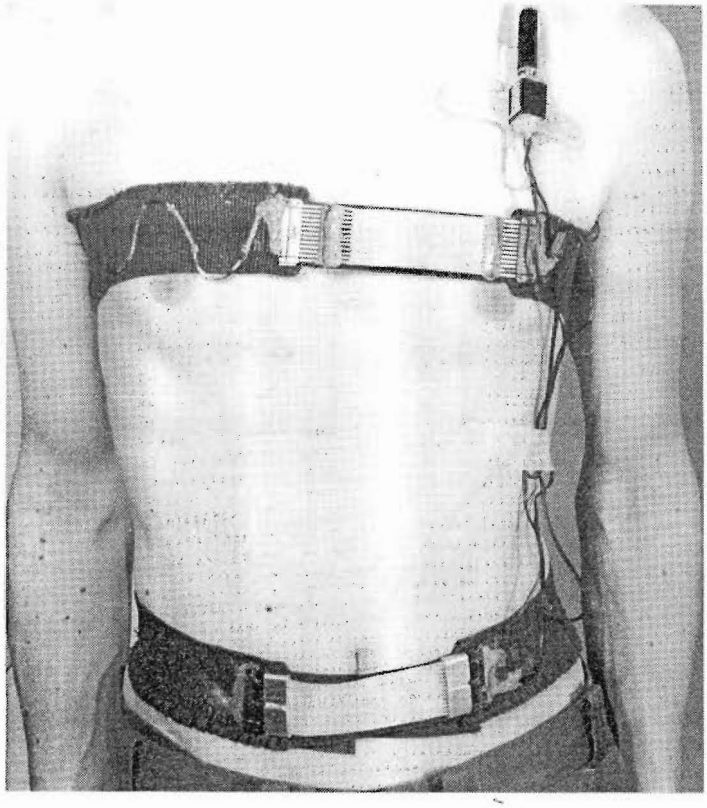


Figure 8.2: Photographs of the respiratory monitoring system.
(a) The complete system.
(b) The inductance transducer belts.

ϕ8.6.1) and stores the sample values onto a floppy disc (Figure 8.1). Typically, the system remains at the bedside overnight, continuously recording the patient's breathing patterns. The sampling process is stopped by entering a command at the keyboard of the VDU, and the system is then disconnected from the patient and taken away from the ward.

The stored data on floppy disc are subsequently analysed (ϕ8.6.2) by software running either on the microcomputer or on a VAX-11/750 minicomputer at the Department of Electrical and Electronic Engineering of the University of Canterbury. The analysis extracts parameters of respiratory performance averaged over intervals called epochs, typically of five minutes duration, and stores the averaged parameters onto disc as processed-data files. Display software running on the VAX-11/750 reads the contents of the processed-data files and generates graphical displays of the averaged parameter values against time-of-day. Long-term variations and trends in respiratory performance of the patient are readily apparent from the displays (ϕ8.7).

The costs of the system hardware in round figures, estimated where necessary, are as follows: inductance plethysmograph $1500 (including labour content for its manufacture), microcomputer $8000, VDU $1500, A/D converter board (ϕ8.5) $500, trolley $500, and ancillary items (e.g. floppy discs, connecting cables) $500. This brings the estimated total cost of the system in late 1983 to NZ$12500.

8.4 THE INDUCTANCE PLETHYSMOGRAPH

The respiratory monitoring system uses an inductance plethysmograph (ϕ7.4.2) as a transducer of lung ventilation. The electronics of the transducer have been adapted from an existing electrical impedance plethysmograph [Seagar 1983a] developed to measure impedance changes in the lower legs during hip replacement surgery [Seagar et al 1984] (see also Chapter 9 which details my subsequent contributions to this work). Inductance plethysmography was initially pursued instead of impedance plethysmography [Hill 1973

ϕ3, Itoh et al 1982a] as a respiratory volume transducer because initial laboratory trials showed it to be more linear than impedance plethysmography, and less affected by motion artefacts from limb movement. Further advantages of the inductance plethysmograph became apparent when it was in use. Firstly, the inductance plethysmograph does not suffer the "heart bump" artefact from cardiac activity that the impedance plethysmograph does [Wilson et al 1982b]. Secondly, to identify abnormal breathing motions such as obstructed or paradoxical motions which may arise from respiratory disease [Sackner et al 1984b] or the administration of analgesic drugs (ϕ8.2.2), it is helpful to make independent measurements of chest and abdomen motions. This is easy to do with the inductance plethysmograph because the two separate coils are electrically coupled only to a comparatively small degree (ϕ8.4.3.4).

8.4.1 TRANSDUCER BELTS

The transducer belts (see Figure 8.3) each consist of a coil of ten turns of 34 SWG enamelled copper wire sewn to an elastic strip in a zigzag fashion to allow the coil to stretch with the elastic. Dimensions are shown in Figure 8.3b. In the clinical setting (i.e. in the hospital wards), particularly with patients recovering from anaesthesia and in pain from their surgical wounds, it is important that the coils be easy to fit to the patient. For this reason the coils are made as a broken loop rather than a closed loop. The elastic is joined at the break by Velcro tapes on its ends, and the electrical circuit is completed across the break by a short piece of flat ribbon cable plugging into flat sockets (Figure 8.3c) at the ends of the zigzag wire. The electrical ends of the coil are brought out to a small two-pin connector (Figure 8.3c) sewn to the elastic of the belt. To improve the electrical and mechanical reliability of the belt, all connections between wires and plug-and-socket connections are covered with silicone rubber window sealant. Short leads run from the chest and abdomen belts to a 6.5mm jack socket mounted on perspex taped to the flat region at the front of the patient's shoulder. The lead to the inductance plethysmograph electronics plugs into the jack socket. This connection is easily disconnected and reconnected if the patient has to leave the bed temporarily.

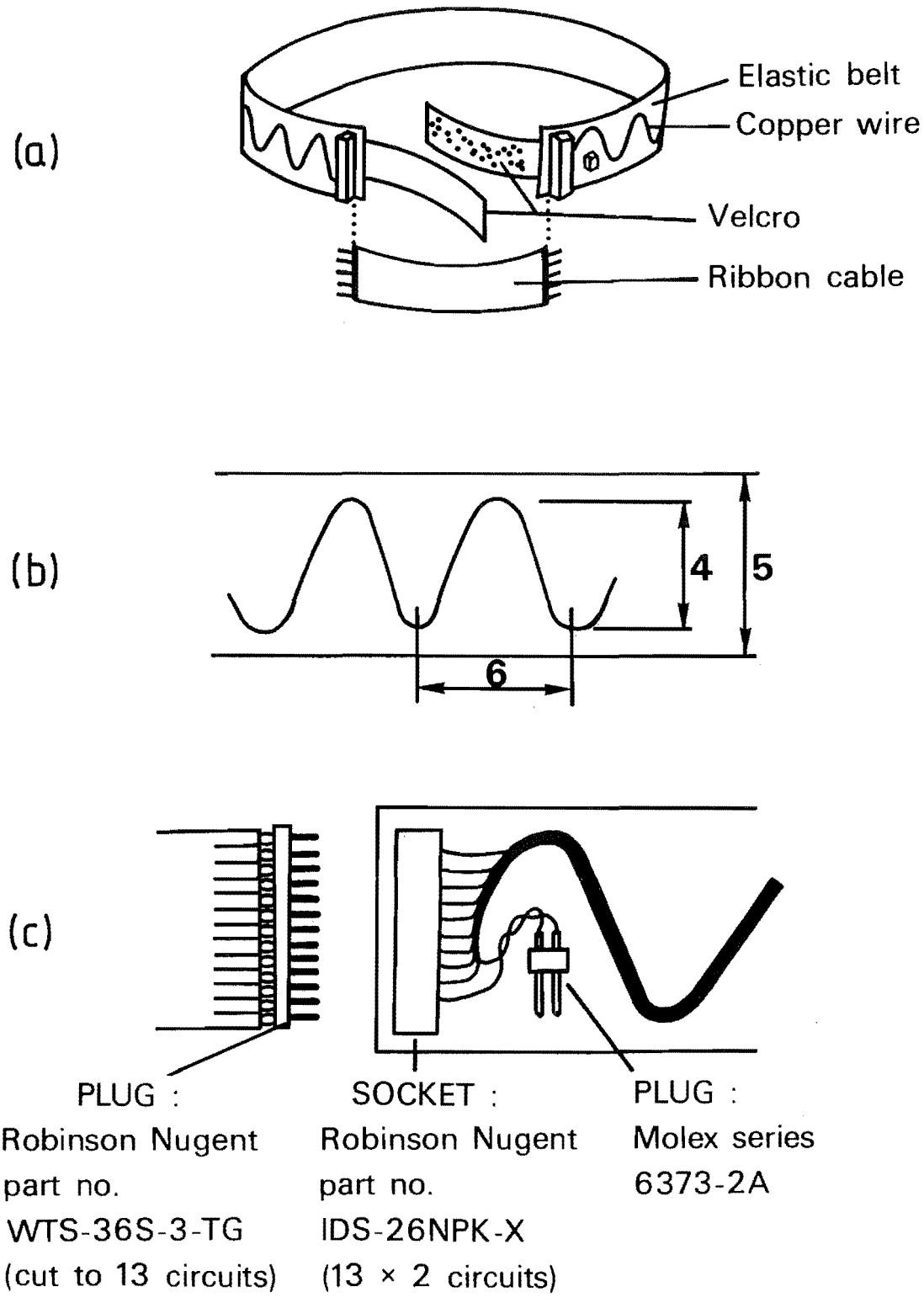


Figure 8.3: Construction of the inductance transducer belts.
(a) Arrangement of belt and ribbon cable joiner.
(b) Dimensions of coil.
(c) Detail of electrical connectors.

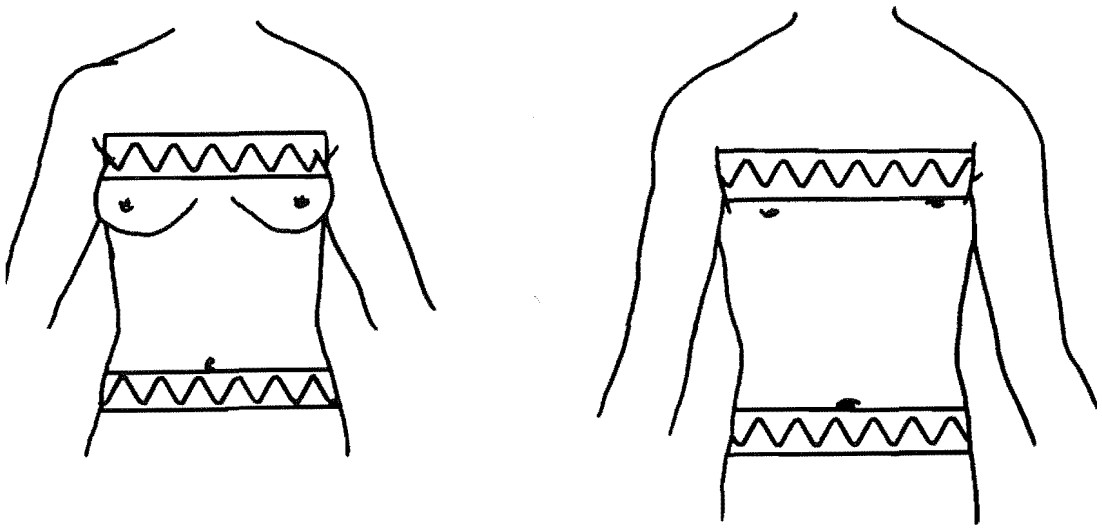
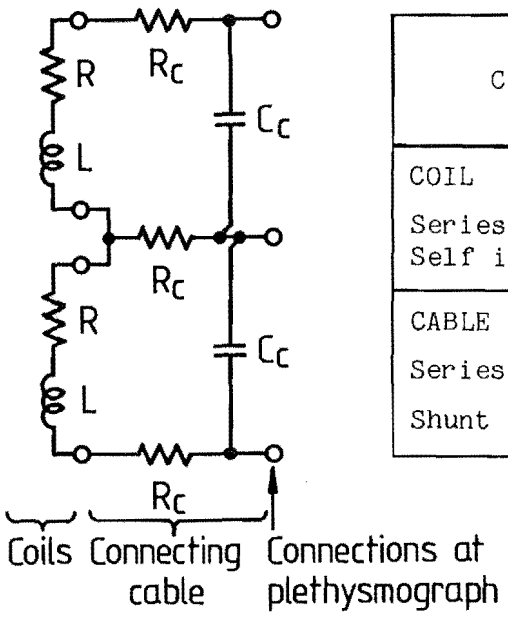


Figure 8.4: Placement of the inductance transducer belts on the patient.



Component	Typical value	Impedance at 62.5kHz (5m cable)
COIL		
Series resistance R	7 Ω	7 Ω
Self inductance L	130 µH	j50 Ω
CABLE		
Series resistance R_c	0.2 Ω/m	1 Ω
Shunt capacitance C_c	300 pF/m	-j1.7 kΩ

Figure 8.5: Electrical arrangement of the transducer coils as used with three-core connecting cable.

The belts are placed on the patient with the chest belt close up under the armpits and the abdomen belt having its top edge over the umbilicus (see Figure 8.4) [cf. Zimmerman et al 1983]. These positions put the chest belt above the breasts in female patients and the abdomen belt between the lower end of the rib cage and the pelvis. Belt placement is altered where necessary to keep the belts clear of the surgical incision. The belts are held firmly in place with surgical tape. It is important that the belts do not move about during the recording, since changing the positions of the coils is likely to alter the calibration of the inductance plethysmograph for respiratory volume.

The magnitude of the inductive impedance of the coils is typically between 40Ω and 70Ω, depending on the size of the patient, for 10-turn coils (see Figure 8.5). Twin-core screened cable, with the screen used as the common signal conductor, connects the patient to the plethysmograph. The series resistance of this cable is about $0.2\Omega.m^{-1}$, and its shunt capacitive admittance has a magnitude of the order of 10^{-4} $mho.m^{-1}$ at 62.5kHz. These values allow the use of cable lengths up to about 100m before the stray impedances begin to dominate the coil impedance that is to be measured. The cable sets currently in use are 5m - 10m in length. The low impedance (less than 100 ohms) of the transducer circuit makes electromagnetic interference unimportant with the inductance plethysmograph in practice, even with the unshielded 10m cables. In comparison, the patient circuit of an impedance plethysmograph can be expected to be more sensitive to electromagnetic interference since the skin-electrode resistance (analogous to the resistance R in Figure 8.5) is usually more than 300Ω for ECG-type electrodes [Beckman Instruments 1965, cf. Hill 1973 ϕ5].

8.4.2 PLETHYSMOGRAPH ELECTRONICS

The inductance plethysmograph measures the impedance of its transducer coils by applying a sinusoidal current of constant 2mA p-p amplitude at 62.5kHz to the chest and abdomen coils in series, and measuring the amplitude of the voltage appearing across each coil. A block diagram of the plethysmograph electronics is shown in Figure

8.6. For each coil (chest and abdomen), a Basal signal and a Filtered signal (with higher gain and a DC block) are made available to the microcomputer. The gains and frequency responses of these outputs are given in Table 8.1. Although not yet implemented in the plethysmograph, provision has been made in the electrical connections to the microcomputer and in the data acquisition software for a "high-frequency" (HF) signal for each coil. This signal is intended to assist with the automatic detection of motion artefacts [Wilson and Franks 1982].

Each voltage receiver consists of a differential-input amplifier with a bandpass response peaking at 62.5kHz and an active rectifier and filter to envelope detect the bandpass-filtered received signal. A DC voltage is produced at the receiver's output proportional to the amplitude of the 62.5kHz AC voltage across its input. For the electrical safety of the patient, the receiver inputs and the constant current source are isolated from electrical ground by transformers.

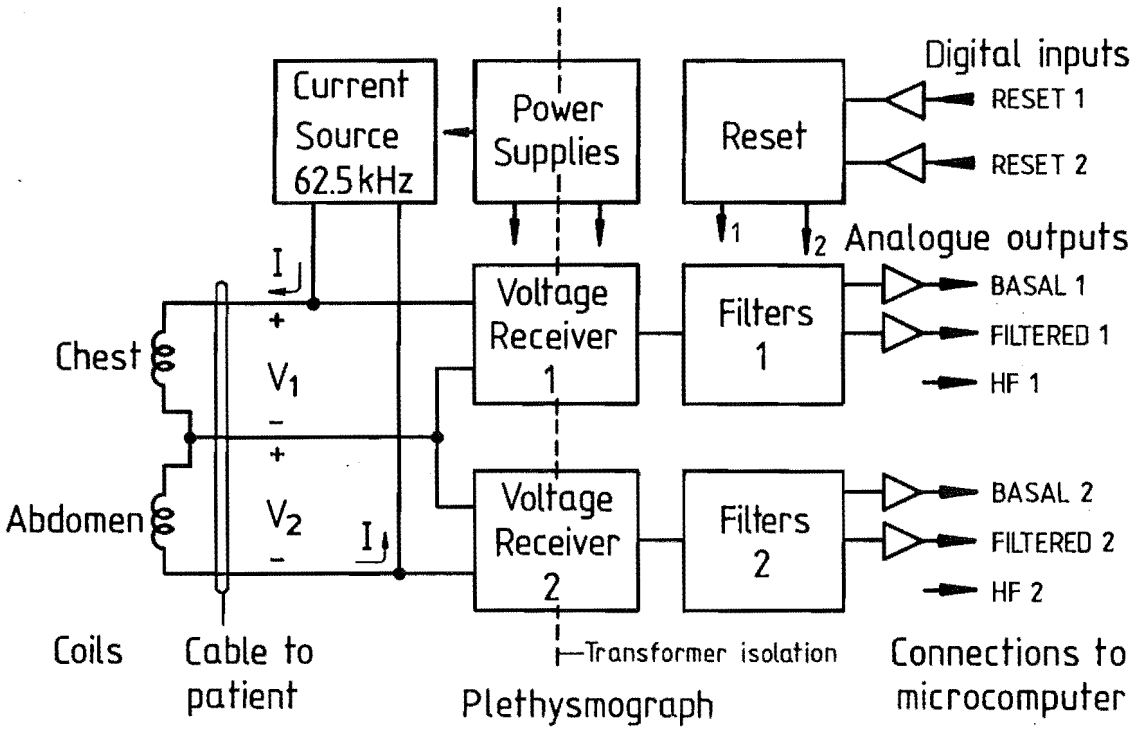


Figure 8.6: Block diagram of the inductance plethysmograph.

The output voltage of each receiver is passed through baseband filters and buffers to provide the analogue outputs appearing at the front and back panel sockets of the plethysmograph. The DC block in each Filtered output comes from a high-pass RC network with a time constant of 10 seconds. A Reset signal from the microcomputer to the plethysmograph allows the Filtered signals to be reset to zero volts under software control if they saturate due to large changes in the basal signals. The analogue outputs are given the mnemonics B1, F1, H1, B2, F2, H2 for the Basal, Filtered, and HF signals of the chest and abdomen respectively.

CONSTANT CURRENT SOURCE	
Isolated from ground	
Short-circuit current	62.5kHz sinusoid, 2mA p-p
Output impedance	> 100kΩ at 62.5kHz
Output voltage	20V p-p maximum
VOLTAGE RECEIVERS (2 channels)	
Isolated from ground, differential inputs, bandpass response around 62.5kHz	
Input impedance	> 50kΩ at 62.5kHz
Input voltage	0.2V p-p nominal maximum for basal full scale
BASAL OUTPUTS (2 channels B1, B2)	
Gain (referred to patient coil impedance)	0.1 V/Ω
Frequency response (-3dB points)	DC - 22 Hz
Normal voltage range	0 - 10 V
FILTERED OUTPUTS (2 channels F1, F2)	
Gain (referred to patient coil impedance)	≈15 V/Ω
Frequency response (-3dB points)	0.016 - 10 Hz
Normal voltage range	-10 - 10 V

Table 8.1: Electrical specifications of the inductance plethysmograph.

8.4.3 PERFORMANCE OF THE PLETHYSMOGRAPH

8.4.3.1 Dependence of Output on Coil Area

The measured relationship between the cross-sectional area of one elastic belt and the basal output voltage of the plethysmograph (which by design is linearly proportional to the inductance of the belt's coil) is shown in Figure 8.7. The area within the belt in these measurements was square, but by virtue of the insensitivity to shape of the coils' inductance (ϕ8.4.3.2), Figure 8.7 can be taken as representative of the performance of the belts on people. The discontinuities in the relationship between area and volume correspond to tucks being let out from the belt to lengthen it as its area was increased. The discontinuities arise because the coil is zigzagged perpendicular to the plane of its cross-section, so that specifying the shape and size of the cross-section does not completely describe

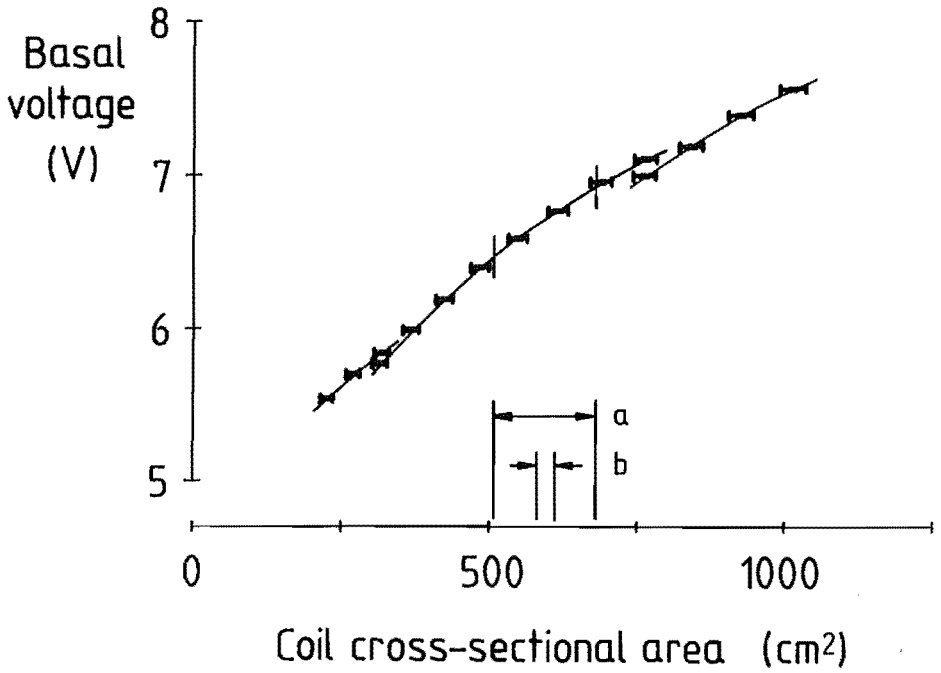
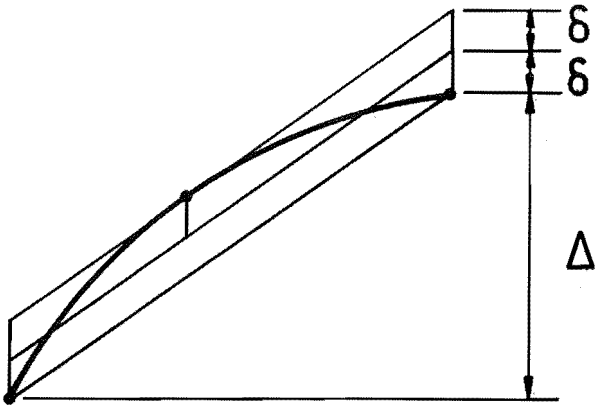


Figure 8.7: Measured relationship between coil cross-sectional area and plethysmograph output voltage, for square cross-sections. The ranges of cross-sectional area marked are for my chest and correspond to (a) vital capacity and (b) tidal volume.

the geometry of the coil.

The range of cross-sectional areas of my chest for the tidal volume of normal breathing and the vital capacity (ϕ7.2) are marked as (a) and (b) respectively in Figure 8.7. The vital capacity range represents the largest possible breathing motion. A quadratic curve fitted by least-squares minimisation to the appropriate measured data in Figure 8.7 deviates from linearity (defined as in Figure 8.8) by ±2.4% over the vital capacity range and by ±0.4% over the tidal volume range.



δ = maximum vertical distance between curve and straight line chosen to minimise δ

$$\text{Nonlinearity} = \pm\frac{\delta}{\Delta}$$

Figure 8.8: Definition of nonlinearity.

8.4.3.2 Dependence of Output on Coil Shape

The basal voltages obtained with the belt in a number of rectangular shapes with the same cross-sectional area but different aspect ratios is shown in Figure 8.9. The aspect ratio is the ratio of the lengths of the longer to the shorter sides of the rectangle. The vertical bars on the measured voltages comprise mostly the

uncertainty in the cross-sectional area of the belt, referred to voltage using the fitted curves of Figure 8.7. The cross-sectional area chosen for these measurements is somewhat smaller than that of the average adult torso (cf. Figure 8.7). Evidently, from Figure 8.9, the output voltage rises slowly with increasing aspect ratio. As in Figure 8.7 there is a discontinuity in the curve where a tuck was let out from the coil to lengthen it as the aspect ratio increased, which shows again the effect of the three-dimensional geometry of the coil.

In Figure 8.9 the relative change of the plethysmograph output voltage between rectangles of aspect ratios 1 and 2 is approximately 0.7%. The aspect ratios of my chest and abdomen, being the ratios of side-side to front-back diameters, are 1.4 and 1.3 respectively. The aspect ratios cannot be less than 1, by definition, and it seems unlikely that aspect ratios of human torsos would ever exceed 2. Therefore, the results of Figure 8.9 indicate that inductance plethysmography can measure cross-sectional areas of the human torso to within ±0.35% accuracy without accurate knowledge of the shape of the cross-sections.

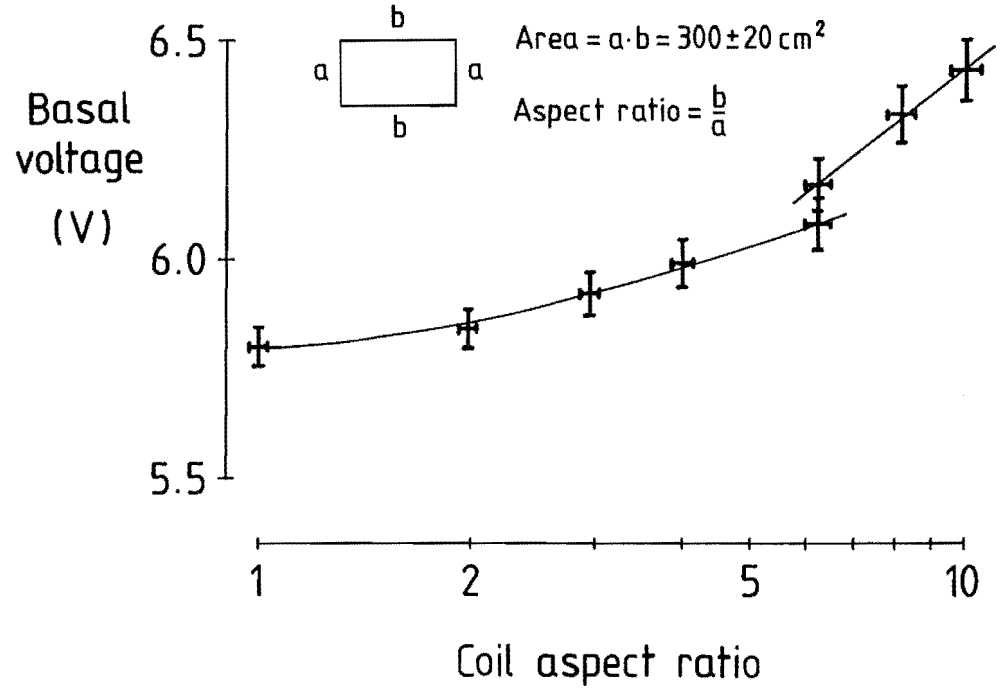


Figure 8.9: Measured relationship between coil shape and plethysmograph output voltage, for rectangular cross-sections of constant area.

8.4.3.3 Effect of Material Within or Close To the Coils

The effect on the plethysmograph output voltage of various materials enclosed by the belt or placed next to it is shown in Figure 8.10. These results are taken from Suntheralingam [1985a]. The belt was fixed around the outside of a round plastic bucket, and various substances were placed in the bucket within the coil. Two sets of measurements were made with different equipment, the first (a - d) using a digital voltmeter to measure the voltage and the second (e - i) using the microcomputer and its A/D converters as a voltmeter. Saline solutions of $5.3g.l^{-1}$ and $42g.l^{-1}$ concentrations gave identical results (c). The steel chair leg (f) was intended to approximate a spinal prosthesis. Measurements (h) and (i) were made with the bucket containing only air, and placed on one of the older-type and one of the newer-type hospital beds in the wards of the Christchurch Hospital in a position approximating that of the torso of a patient lying in the bed.

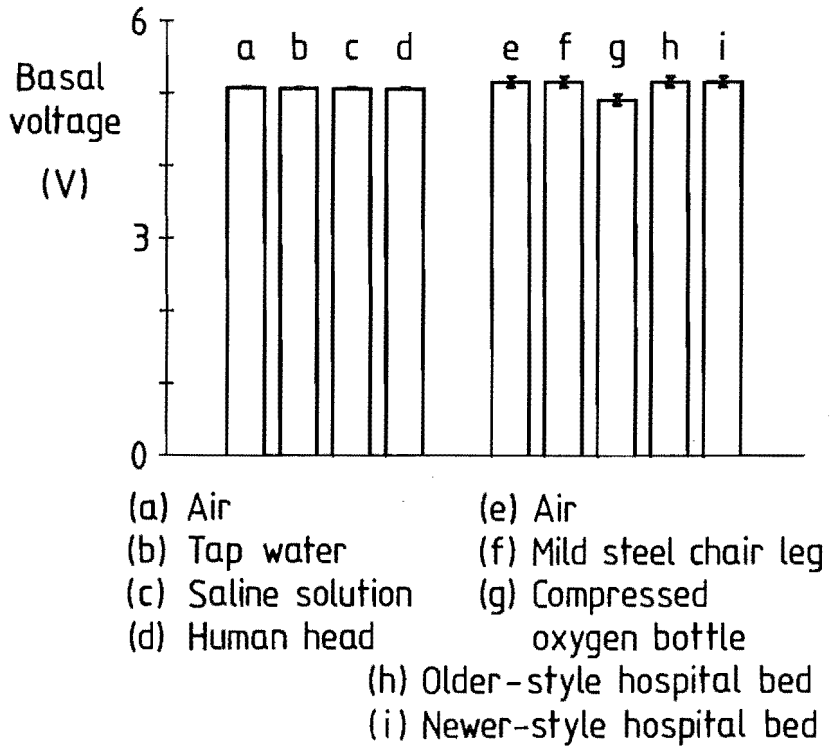


Figure 8.10: Measured effect of placing different materials within (a - g) or near to (h, i) the plethysmograph coil.

Of all the materials placed within or near the belt, the only one which produced a significant change in the inductance of the coil was a 335 litre oxygen bottle, measurement (g). This shows that the inductance of the coils is unlikely to be significantly affected by the composition of the patient or the coils' proximity to the steel wire in the hospital beds. These factors therefore need not be of concern when the inductance plethysmograph is used for quantitative measurements of respiratory volume.

8.4.3.4 Mutual Coupling Between the Coils

The two coils of the respiratory inductance plethysmograph act as two windings of a transformer, because they are loosely coupled magnetically. The magnetic coupling of each coil to itself and to the other coil is described by the self inductances L_1, L_2 and the mutual inductance M (see Figure 8.11) [Skilling 1974 ϕ11]. Since M depends on the geometry of both coils, the voltage induced in each coil is partly dependent on the geometry of the other coil. This produces electrical coupling between the chest and abdomen voltage outputs of the respiratory inductance plethysmograph, additional to whatever mechanical coupling of the body motions may exist. Capacitance between the two coils and between the coils and the patient introduces further electrical coupling between them.

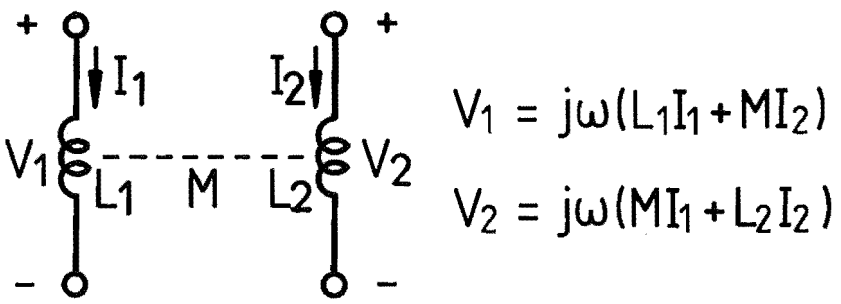


Figure 8.11: Magnetic coupling of the two coils of the inductance plethysmograph, and definitions of self-inductance L_1 and L_2 and mutual inductance M for sinusoidal currents I_1 and I_2 at angular frequency ω.

Figure 8.12 shows the results, as a function of distance d between the coils, of direct measurements of the amount of coupling between the plethysmograph's output voltages. The measurements were obtained with the geometry of Figure 8.13 by perturbing the area of

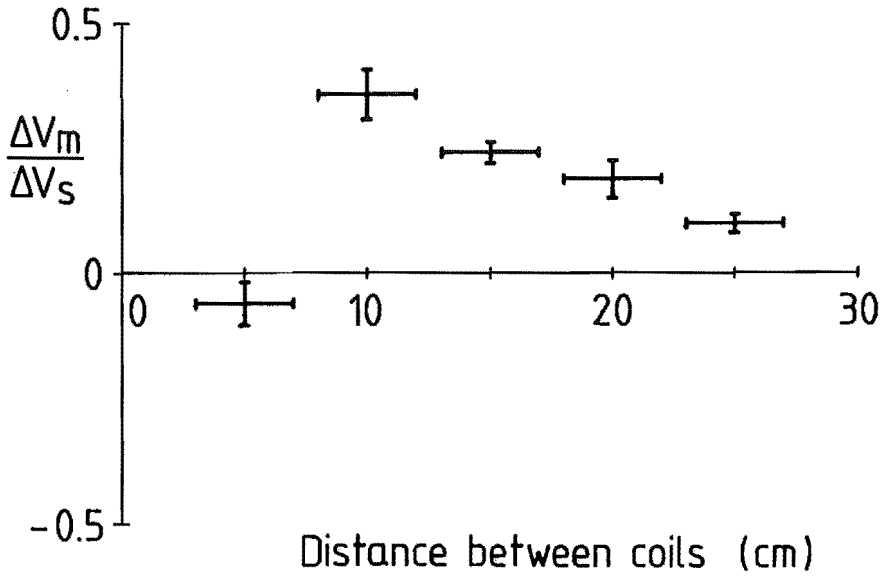


Figure 8.12: Ratio of mutual voltage change ΔV_m to self voltage change ΔV_s (see text) as a function of distance d between the plethysmograph coils (see Figure 8.13).

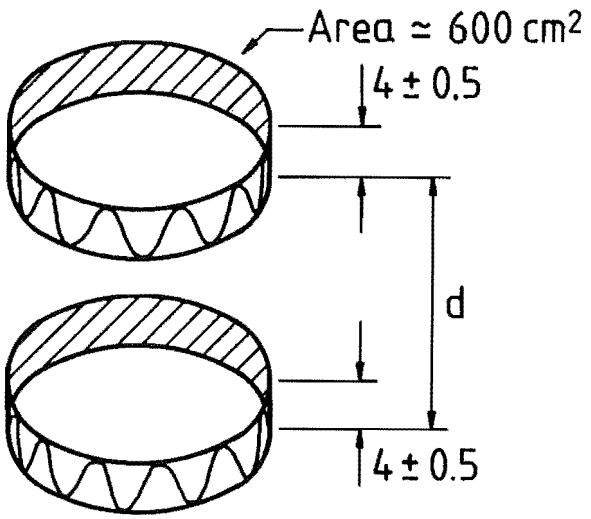


Figure 8.13: Geometry of plethysmograph coils for measurement of mutual coupling effects (see text). The coils were worn on a human torso. Unmarked dimensions are in cm.

one belt and measuring the self voltage change ΔV_s in that coil and the mutual voltage change ΔV_m in the other coil. The belts were worn by a person, so that capacitance effects between the coils and the person were present. The area of one coil was perturbed by lifting a part of it away from the person's body.

For distances d ≥ 10cm the mutual inductance of the coils was the dominant coupling mechanism. The ratio $\Delta V_m/\Delta V_s$ was positive and decreased with increasing d. For d = 5cm (note from Figure 8.13 that the bottom of one coil was then only 1cm from the top of the other coil) the capacitance between the coils was an important part of the electrical coupling. The ratio $\Delta V_m/\Delta V_s$ was negative, presumably due to the capacitance between the coils decreasing as one coil was lifted away from the other. This situation should not occur in practice since the belts remain close to the body surface and cannot be perturbed independently.

When the respiratory inductance plethysmograph is in normal use (d ≥ 15cm), the ratio $\Delta V_m/\Delta V_s$ which expresses the amount of coupling between the chest and abdomen is likely to be less than 25% (Figure 8.12). It may be greater than this for obese patients, on whom the diameter of the coils is larger relative to the distance between them. Provided the belts remain in fixed positions on the patient so that their mutual coupling does not change during calibration and measurement, the mutual coupling can be compensated for by the calibration procedure (ϕ7.4.4) and so should not impair the quantitative measurement of respiratory volume by inductance plethysmography.

8.4.3.5 Measurement of Respiratory Volume and Flow

The performance of the respiratory inductance plethysmograph as a transducer of respiratory volume and flow is illustrated in Figure 8.14, in which volume and flow signals derived from the Filtered outputs of the inductance plethysmograph are compared with volume and flow signals obtained simultaneously from an Ohio Model 840 spirometer [Tisi 1980 ϕ4]. In this example the signals were sampled at 0.5 second intervals, and the impulse response of the DC blocking

network in the Filtered channels (ϕ8.4.2) was deconvolved from the Filtered signals by the computer. The calibration coefficients for the inductance plethysmograph (ϕ7.4.3) in this example were obtained manually, using the isovolume part of the recordings (the interval in

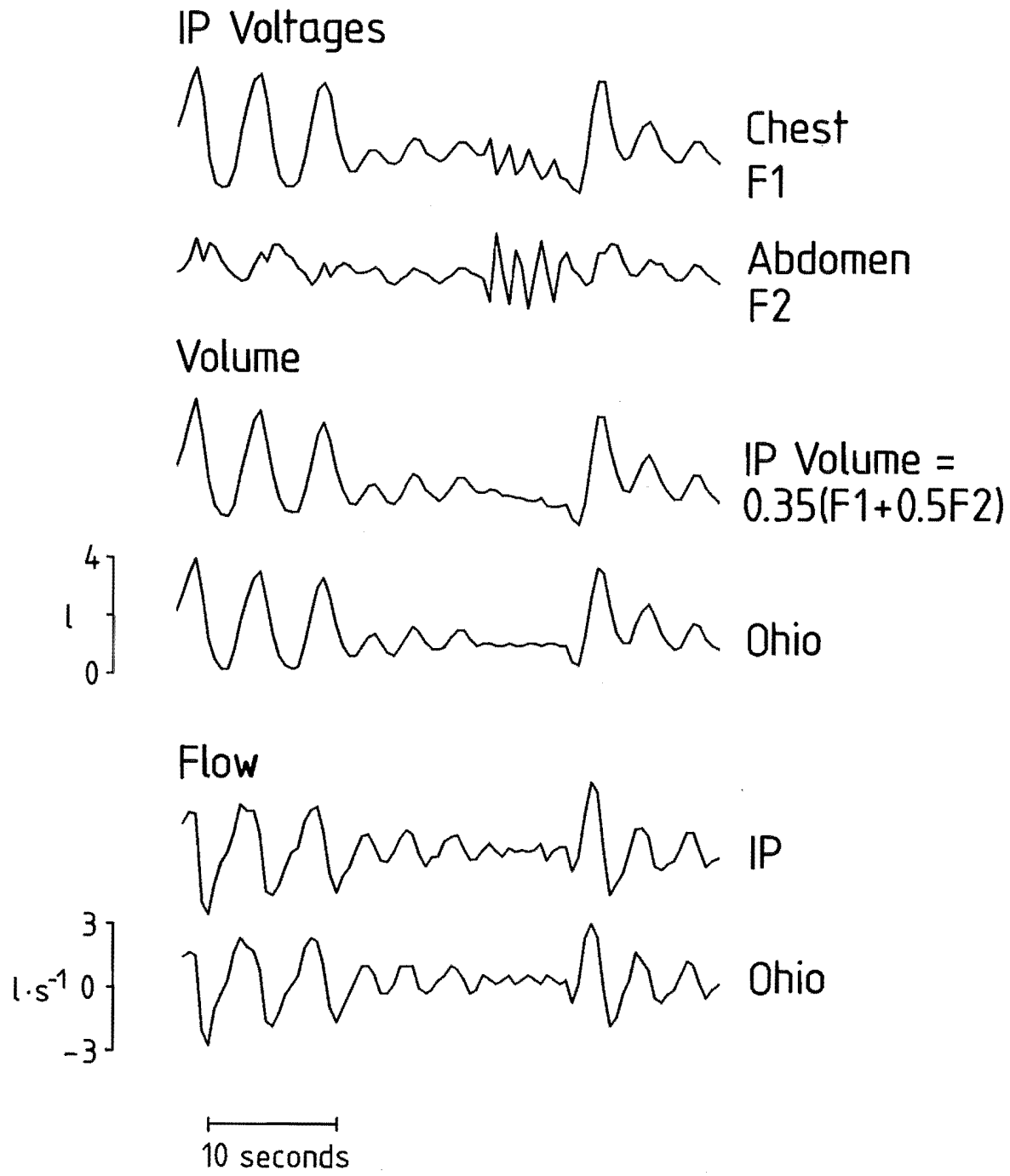


Figure 8.14: Volume and flow signals from the inductance plethysmograph (IP) compared with those from an Ohio model 840 spirometer.

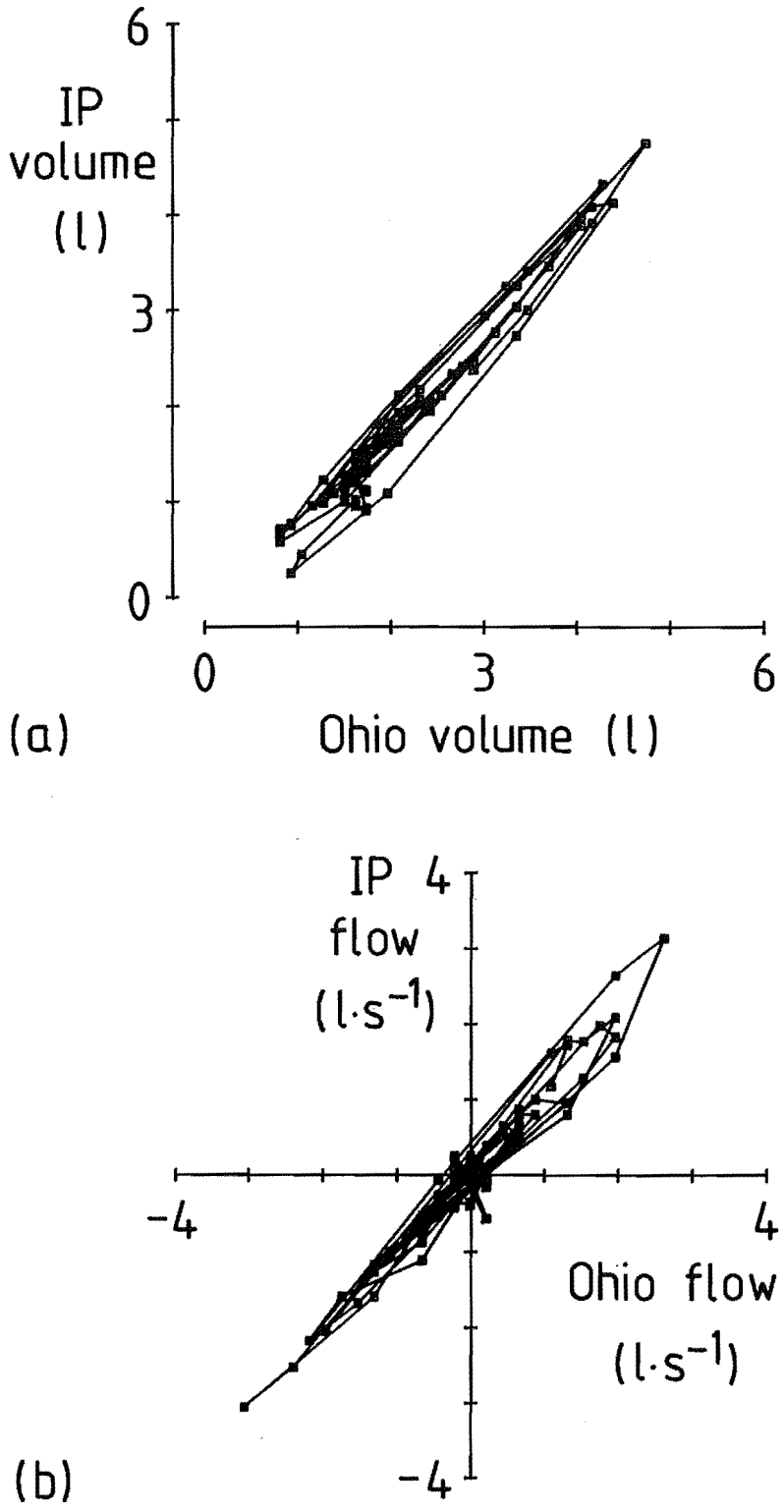


Figure 8.15: X/Y plots of
(a) inductance plethysmograph (IP) volume against Ohio spirometer volume,
(b) inductance plethysmograph flow against Ohio spirometer flow,
for the signals of Figure 8.14.

Figure 8.14 when the volume and flow signals are nearly constant) to obtain the relative weights for the plethysmograph voltages and then the whole recording to obtain the absolute values of the weights. The flow signals were obtained from the first differences of the respective volume signals.

The spirometer volume signal and the plethysmograph volume signal are closely similar in Figure 8.14. The spirometer and plethysmograph flow signals compare less well, but are still quite similar to each other. The subjective similarity of these curves demonstrates the capability of the inductance plethysmograph to faithfully measure lung volume changes, provided the plethysmograph is properly calibrated.

The spirometer and plethysmograph volume and flow signals from Figure 8.14 are plotted against each other for volume in Figure 8.15a and for flow in Figure 8.15b. The value of Pearson's product-moment correlation coefficient [Sokal and Rohlf 1981 ϕ15.2] between the spirometer and the plethysmograph signals is 0.982 for volume and 0.978 for flow. Hill et al [1982] performed the same tests on a respiratory inductance plethysmograph similar to the commercially produced Respitrace equipment [Watson 1979, Sackner et al 1980b], and obtained somewhat better correlation between the spirometer and plethysmograph than shown here. The poorer correlation here is probably due to the variety of breathing manoeuvres present in the recordings of Figure 8.15, compared to the steady breathing used by Hill et al [1982].

8.4.3.6 Example of a Clinical Recording

Figure 8.16 shows excerpts from a recording made at the Christchurch Hospital. The subject was a male of 49 years age and 87kg weight, and the recording was made before surgery so his breathing was "normal". The two signals shown are the Filtered signals for chest and abdomen, the samples were taken at one second intervals, and the timing ticks are at one minute intervals. The first excerpt (Figure 8.16a), from the mid evening when the patient was awake, shows the severe artefacts in the recording from gross body

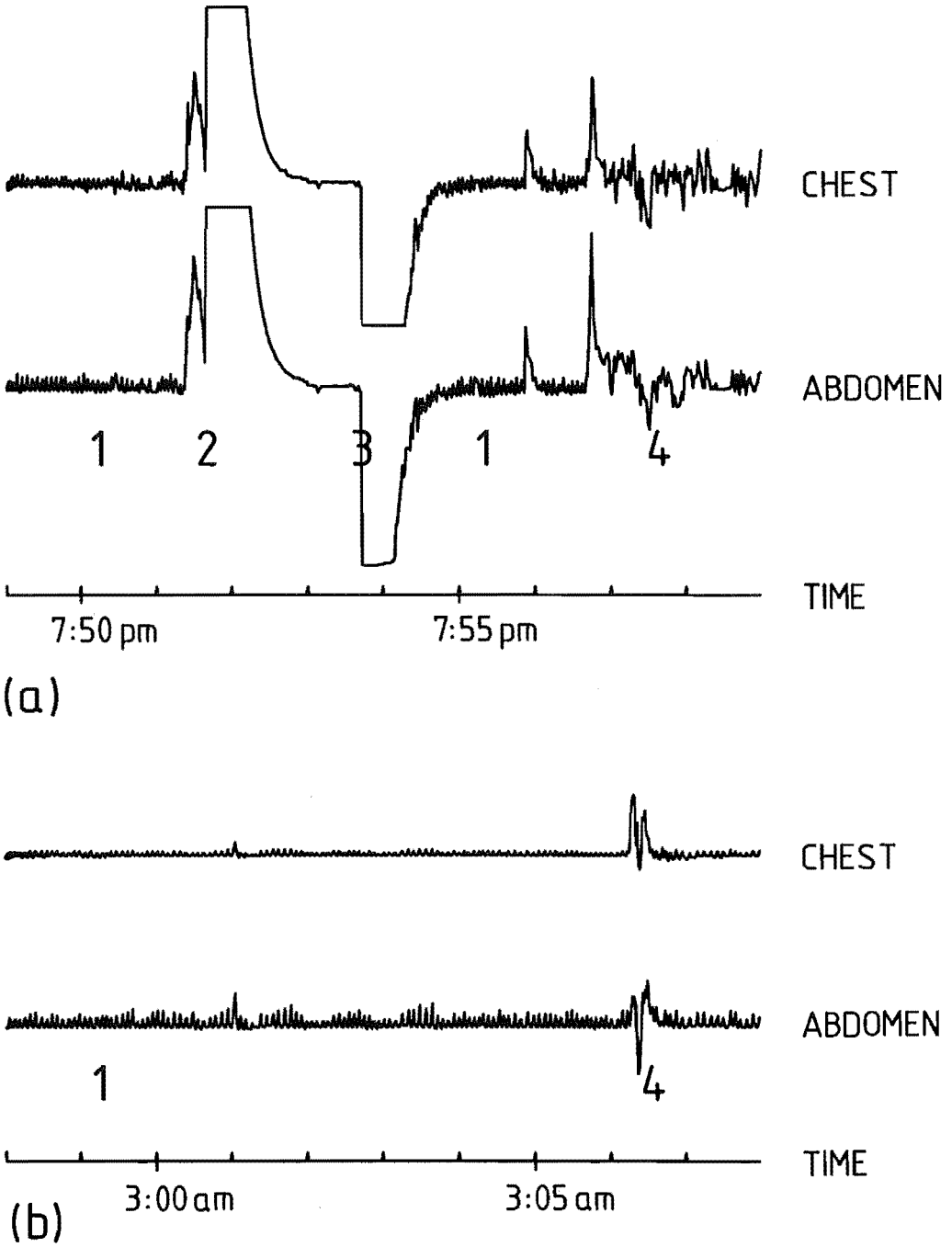


Figure 8.16: Excerpts from a clinical recording of a patient's breathing motions, before surgery. The timing ticks show one minute intervals.

(a) Recording from mid evening when the patient was awake.

(b) Recording from early morning when the patient was asleep.

The following features are present:

(1) Normal breathing, (2) disconnection from the plethysmograph, (3) reconnection to the plethysmograph, (4) motion artefacts.

motion and disconnection of the patient from the equipment. The low amplitude oscillatory segments are normal breathing. The second excerpt (Figure 8.16b) from the early morning when the patient was asleep, is typical of the much better quality of recording obtained when the patient is sleeping and moves only occasionally. Differences in breathing patterns between wakefulness and sleep can be seen in these excerpts, notably a decreased relative amplitude of the chest motions and a periodic variation of amplitude (cf. periodic breathing [Best and Taylor 1972 ø6]) in the early morning recording.

8.5 THE MICROCOMPUTER

The task for which the respiratory monitoring system is designed is to detect correlations between analgesia and changes in respiratory patterns, so that the respiratory effects of different types of analgesia can be compared. To detect these correlations in the presence of the considerable random variability of normal respiratory waveforms, it will be necessary to statistically analyse a great many breaths. Each patient is to be monitored for perhaps twelve to twenty-four hours before and after surgery, and it may prove necessary to monitor hundreds of patients to obtain statistically significant results. Enormous volumes of data will be generated, and automated analysis is necessary for the analysis of this data to be practical. A microcomputer has therefore been included in the respiratory monitoring system to automate data collection by the system, to perform as much data analysis as possible in real-time, to perform any remaining data processing retroactively, and to allow access to and display of the processed results.

The microcomputer is a general-purpose machine of the "small business computer" class, and is capable of useful tasks apart from its function as part of the respiratory monitoring system. As an important example of this, a large part of the software for the respiratory monitoring system, including all of the data acquisition software (ø8.6.1), has been developed on the microcomputer itself.

8.5.1 HARDWARE

The computer is a Systems Group model 2812 microcomputer [Systems Group 1981]. It has a Z80-A CPU running at 4MHz, 64K of dynamic RAM, twin Shugart 801 8" floppy disc drives providing single- or double-density formats at 608K bytes per drive in double density, and is being used with the CP/M 2.2 operating system. The machine uses the IEEE standard S-100 bus internally [Elmquist et al 1979, Stewart 1983] and has a number of spare slots for insertion of custom PC boards for special functions. It provides four independent serial I/O channels, one of which is used with a standard VDU as the normal means of communication between the computer and its operator.

Specific hardware support for the respiratory monitoring function is provided by a Z8430A counter-timer chip on the CPU board and by an analogue/digital input/output board purchased separately and installed into the computer. The timer chip provides real-time interrupts by which the timing intervals between samples of the inductance plethysmograph signals are accurately controlled. The analogue/digital input/output board provides seven channels of 8-bit analogue-to-digital (A/D) conversion, seven channels of 8-bit digital-to-analogue (D/A) conversion, and a bidirectional 8-bit parallel digital port. These signals are brought out to the rear panel of the microcomputer with suitable signal buffering for the connection by cable to the inductance plethysmograph. The analogue inputs receive the six voltage signals (ϕ8.4.2) from the plethysmograph. The analogue outputs are used to drive chart plotters for generating copies on paper of the recorded breath waveforms. The digital output lines are used as control signals to the plethysmograph, for example the Reset signal to each plethysmograph channel (Figure 8.6).

8.5.2 SOFTWARE DEVELOPMENT AND OPERATING ENVIRONMENT

The software which implements the data acquisition and processing algorithms on the microcomputer has been written in Pascal/MT+ [Digital Research 1981], with a small amount of Z80 assembly language included in the low-level routines for programming

and driving the sample timing hardware. The Pascal/MT+ package [Digital Research 1982] provides an excellent environment for Pascal software development. The screen editor provided with the package gives access within the editor to fast program syntax checking, which greatly speeds the creation of correct programs. The compiler and linker support modular compilation, in which several parts of a program may be created as separate "modules" and compiled independently then linked together subsequently to produce the working program. Modular compilation encourages structured programming, i.e. a careful hierarchical decomposition of the problem and the software solution to it [Bergland 1981], and speeds program development since any given change requires recompilation of a single module, not the entire program. It is particularly important to eliminate unnecessary compilation when developing software on a microcomputer, because the limited computational power makes compilation take a relatively larger part of the time of the edit/compile/run cycle of software development than it does on a more powerful machine such as a minicomputer.

By carefully defining the program structure and the function of the modules, several programs whose functions are related but different overall are able to use common modules that perform common tasks. The respiratory monitoring software has been written using this approach. Each program has the same main module which is linked to other modules providing generic functions such as reading a sample, processing a sample, and timer control. Different versions of the modules for each generic function are written and one version for each generic function is chosen when linking a program together. Many different program functions, such as reading samples from plethysmograph to disc, reading samples from disc to chart plotter, breath processing from disc, or real-time breath processing from the plethysmograph, are possible by appropriate selection from the basis set of modules. This approach carries the overhead of extra care and restrictions during the design of the main program and each module, but overall it eases the burden of developing and maintaining several programs having several related functions.

The respiratory monitoring software executes under the CP/M 2.2 operating system [Zaks 1980]. CP/M appears to the software as a collection of facilities for disc file handling and communication with

the VDU. In this role CP/M 2.2 is adequate to the requirements of the respiratory monitoring software. To the user of the microcomputer CP/M appears as a prompt character on the VDU screen, to which replies (commands) are typed at the VDU keyboard. The clinical staff who use the system for data collection have little contact with CP/M, because after turning the computer on and inserting a floppy disc into it they need enter only a single CP/M command to invoke the respiratory monitoring program.

8.6 DATA PROCESSING

The data processing performed by the microcomputer encompasses data acquisition, or reading and storing the samples, and data analysis. Analysis of the data, to characterise the respiratory effects of analgesia, has two goals. The first is to accumulate statistical information describing the breathing, and to identify and display long-term trends in respiratory performance. The second is to detect in real-time any short-term irregularities of the breath waveform, e.g. apnoeic pauses or paradoxical breathing [Jordan 1982], so that detailed recordings of the respiratory waveform surrounding anomalies can be made for later examination by clinicians. The data processing of the respiratory monitoring system currently implements the first of these two goals.

The data analysis algorithms have been deliberately kept simple so that they can be performed in real-time by the respiratory monitoring system's microcomputer, although the data analysis is at present performed only after the data acquisition is complete. Real-time data processing introduces important improvements in the capability of the system. These are discussed in Section 8.8.

8.6.1 DATA ACQUISITION

Breath data are acquired by the software in the form of "sample records". Each sample record includes six integer numbers containing the values of the six analogue output voltages of the inductance plethysmograph (ø8.4.2) at the instant the sample was

taken. The interval between samples is established by real-time interrupts generated every 8ms by the Z8430A counter-timer chip on the microcomputer's CPU board. The software counts the number of interrupts since the last sample and takes a new sample when the chosen sample interval has elapsed. When each sample is read, it is written to the VDU screen in a graphical format to provide an immediate display of the inductance plethysmograph output voltages, and is also stored in a data file on floppy disc for later retrieval by the data analysis software. The data file is created at the beginning of the monitoring session, and the data in it begin with a file header containing information such as the date and time of the recording and the sampling interval.

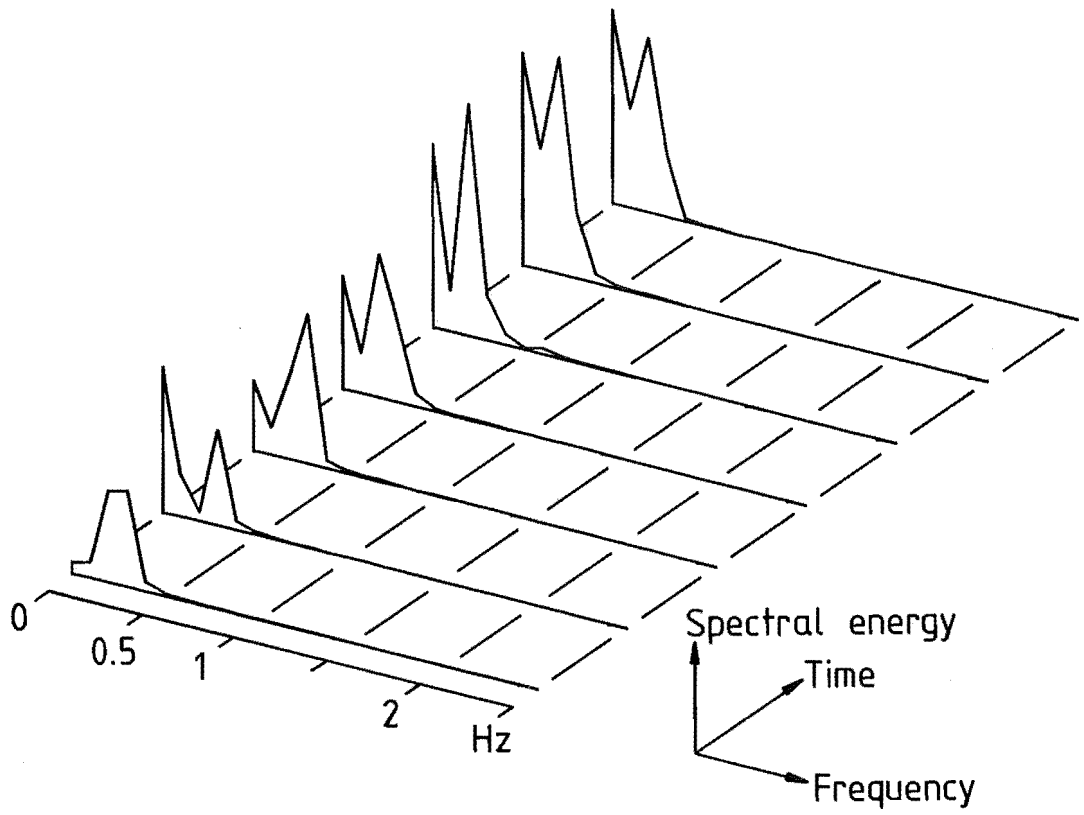


Figure 8.17: Fourier spectrogram of normal breathing, which was sampled at 80ms intervals and fast Fourier transformed in contiguous records of approximately 13 seconds each. The two peaks in most of the spectra are the "fundamental" breathing frequency at 0.2 - 0.3Hz and the DC and low frequencies below the fundamental.

The respiratory monitoring system is used in the wards with a single double-density floppy disc containing only the data acquisition software and essential CP/M utility programs. The disc has about 550K bytes of free space, which can hold about 80,000 samples since each sample takes seven bytes in the data file. Sampling once per second gives about 22 hours' recording on the one disc. Spectral analysis of normal breathing sampled at 80ms intervals (see Figure 8.17) shows that the spectral energy above the "fundamental" peak in the breath spectrum is small. In Figure 8.17 the breathing rate is between 0.2Hz and 0.3Hz, and less than 4% of the total spectral energy is at 0.5Hz and above which is the frequency range for which the requirements of the sampling theorem [Bracewell 1978 ϕ10] are not met by sampling once per second. This suggests that aliasing should not be a cause of severe difficulties provided the breathing period is appreciably greater than two seconds, which it usually is in adults at rest [Guyton 1971 ϕ29]. The limited temporal resolution of one second sampling is evident in the TIBB frequency stackplots presented in Sections 8.6.3 and 8.7.

8.6.2 DATA ANALYSIS

The data analysis software operates on a sequential stream of samples, which are read from the data file created during data acquisition (ϕ8.6.1). For real-time data analysis the sequential stream can equally well come directly from the inductance plethysmograph. The analysis has the following steps, which are described more fully in the following subsections. The Filtered signals for chest and abdomen in the sample records are scaled and added to produce a lung volume signal. Individual breaths are detected from this volume signal. The length (or TIBB, Time Interval Between Breaths) and amplitude (or tidal volume) of each breath are found. The TIBB and tidal volume information are averaged over five-minute epochs and written to a processed-data file at the end of each epoch. Minute ventilation is computed from the tidal volumes and stored for each epoch, and also a phase index describing the relative phase of the chest and abdomen signals is averaged and stored for each epoch. Reading the processed-data file and displaying the processed values for each epoch against time produces displays (ϕ8.6.3) on

which long-term trends or patterns of change in the parameters are readily apparent to the eye.

In the following exposition n is the index of individual samples read during the epoch and m is the index of individual breaths detected during the epoch. There are N samples read and M breaths detected in an epoch, whose duration is T. M and possibly N vary from epoch to epoch. The length of the mth breath is τ_m.

8.6.2.1 Scaling to Volume

The data analysis software uses the Filtered signals for the chest and abdomen, which are high gain versions of the impedances of the chest and abdomen transducer belts with the DC component or basal impedance removed. These signals appear to the program as numbers with a certain scale factor relative to the output voltages of the inductance plethysmograph, and the plethysmograph voltages themselves are scaled versions of the volume changes of the chest and abdomen compartments (ϕ7.4.3) from their average values (cf. 7.1, 7.4). To scale the chest and abdomen numbers received by the program into volumes, in litres say, the constants a and b in (7.4) must be found by a calibration procedure such as one of those described in Section 7.4.4.

Since quantitative calibration for volume has not yet been implemented for the respiratory monitoring system, the data analysis software currently uses arbitrary values of unity for the scale factors of both the chest and abdomen signals.

8.6.2.2 Detection of Breaths

Breaths are detected in the incoming sample stream by identifying the onset of inspiration and expiration and defining a breath as the set of samples from the beginning of one inspiration to just before the beginning of the next. The onset of inspiration and expiration is found simply by examining the sign of the first

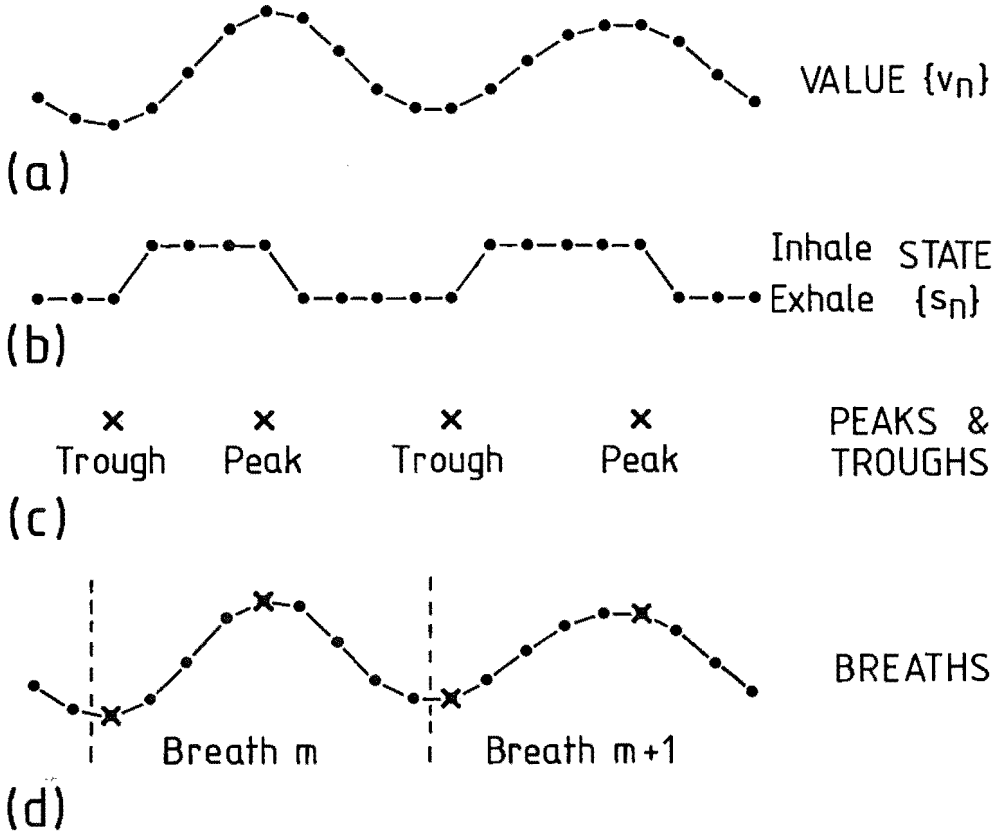


Figure 8.18: Detection of breaths in a sampled signal. $\{v_n\}$ is the sampled signal's value (a) and $\{s_n\}$ is its state (b). Changes in state occur at peaks and troughs (c), from which breaths are defined (d).

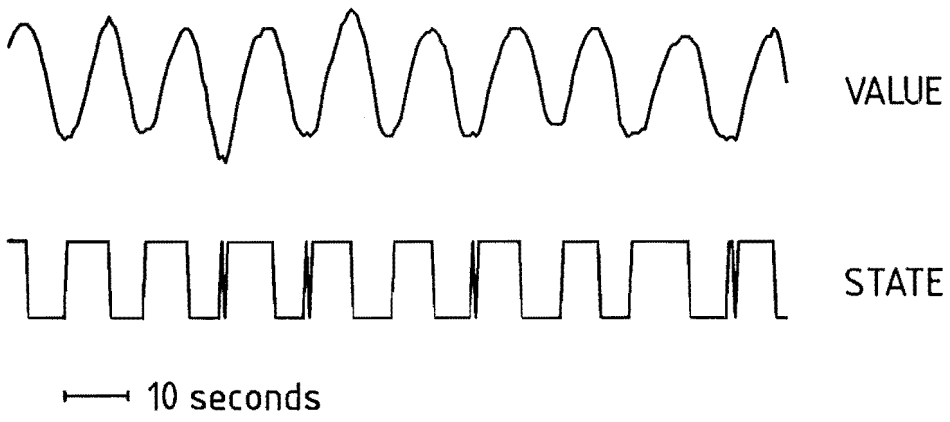


Figure 8.19: Illustrating the performance of the breath detection algorithm on a signal of good quality.

differences of the incoming sample values. The detection of breaths in a particular signal (e.g. chest volume, abdomen volume, or total volume) obtained from the incoming samples is illustrated in Figure 8.18.

The samples $\{v_n\}$ of the signal's value (Figure 8.18a) are numeric values which may be compared to one another and be found to be less than, equal to, or greater than each other. The first step in breath detection is to determine the state $\{s_n\}$ of the signal which signifies whether the signal represents inhalation or exhalation (Figure 8.18b). This is done according to the rules

$$\begin{aligned} v_n > v_{n-1} &\Rightarrow s_n = \text{INHALE}, \\ v_n < v_{n-1} &\Rightarrow s_n = \text{EXHALE}, \\ v_n = v_{n-1} &\Rightarrow s_n = s_{n-1}. \end{aligned} \tag{8.1}$$

The states of inhalation and exhalation are conveniently represented within the computer program as the numeric values +1 and -1 respectively, i.e.

$$\begin{aligned} \text{INHALE} &= +1, \\ \text{EXHALE} &= -1 \end{aligned} \tag{8.2}$$

in (8.1).

From the state values $\{s_n\}$ the peaks and troughs of the signal are found (Figure 8.18c) by the rules

$$\begin{aligned} s_{n-1} = \text{EXHALE and } s_n = \text{INHALE} &\Rightarrow \text{trough at } n-1 \\ s_{n-1} = \text{INHALE and } s_n = \text{EXHALE} &\Rightarrow \text{peak at } n-1 \\ s_{n-1} = s_n &\Rightarrow \text{no peak or trough at } n-1. \end{aligned} \tag{8.3}$$

The peaks and troughs are easily found directly from $\{v_n\}$, but it is useful to generate $\{s_n\}$ as an intermediate step because $\{s_n\}$ is later used directly in computing the chest/abdomen phase index (ø8.6.2.7). Each trough in the incoming signal is taken as the start of a new breath (Figure 8.18d).

The breath detection scheme described above is rudimentary and

sensitive to noise on the respiratory signals, but it has been quite effective in practice, aided by the simple artefact detection algorithm of Section 8.6.2.3, except when the patient is bodily moving about. The quantisation of the signal in time (by the sample interval) and value (by the resolution of the A/D conversion) discriminates against high-frequency low-amplitude noise. In this sense it is desirable that the sampling interval be as long as possible consistent with satisfying the requirements of the sampling theorem [Bracewell 1978 ø10] relative to the periods of the breathing (there must be at least two samples per cycle of breathing otherwise a peak and a trough cannot be detected for each breath). The sample interval of one second used in the clinical recordings made so far seems to be a good compromise between fidelity of representation of the signal and noise immunity. A small amount of noise on the signal causes the breath detection to report the presence of spurious, short breaths (see Figure 8.19, which was sampled at 0.5 second intervals). The effects of large amounts of noise and the quantisation of the TIBB measurements by the sample interval can be seen in the TIBB frequency stackplots presented in Section 8.6.3.2.

Detection of breaths in the presence of high-frequency low-amplitude noise can be made more reliable by temporally filtering the sampled signal [cf. Spencer et al 1972, Sackner et al 1984b] and by requiring the amplitude of oscillations accepted as breaths to exceed some small threshold value, perhaps determined from the amplitudes of previously detected breaths [cf. Wilson and Franks 1982]. The potential efficacy of these measures is clear from Figure 8.19, in which the spurious short "breaths" detected at the peaks and troughs of some of the true breaths could be easily suppressed. The need for explicit rejection of high-frequency low-amplitude noise in the signal becomes greater as the sample rate is increased above once per second. The sample rates of the computerised respiration analysis systems described in the three references quoted immediately above are all between 20 and 30 samples per second.

Wilson et al [1982a] briefly compare peak and trough detection with zero-crossing detection (i.e. locating the instants at which the oscillatory waveform crosses a threshold representing its average value over, say, its last few cycles) for detecting breaths in

respiratory signals. They conclude that peak and trough detection is more suitable for detecting breaths, in part because zero-crossing detectors more easily miss breaths if the baseline value of the respiratory signal changes suddenly.

8.6.2.3 Detection of Artefacts

When a "breath" is detected, tests are applied to determine whether it is an artefact from noisy or corrupted data. The first test classifies a "breath" as an artefact if either of the Basal voltages (ϕ8.4.2) exceed the nominal maximum of +10V within the duration of the breath. This occurs if the connections to the patient become open circuit, as can happen routinely when the patient is unplugged from the plethysmograph. The second test classifies a "breath" as an artefact if its amplitude (i.e. tidal volume) exceeds a threshold value, which in the present system has a fixed value specified by the operator when the data analysis program is run. This test detects motion artefacts, whose amplitudes are usually much greater than the amplitude of quiet breathing (Figure 8.16). A "breath" which is classified as an artefact is not used in the computation of the breathing parameters described in the following subsections.

More sophisticated tests for noise and artefacts in the signal are readily devised. Wilson and Franks [1982] report that the high-frequency content of motion artefacts is greater relative to the total spectral energy than that for normal breathing, and they include special hardware to provide this information in their respiratory monitoring system. The HF channels mentioned in Section 8.4.2 are an advance provision for the same hardware in the system described here, following Wilson and Franks [1982]. Ad-hoc schemes for rejecting particular types of artefact may also be implemented [e.g. Wilson et al 1982a].

8.6.2.4 Time Interval Between Breaths

The Time Interval Between Breaths (TIBB) τ is defined as the

interval between the beginnings of two successive breaths (see Figure 8.20). The software maintains a counter which is incremented as each new sample is received, and computes the TIBB from the number of samples between successive troughs multiplied by the sample interval which is read from the data file header (ϕ8.6.1).

The mean and standard deviation of the TIBB values τ_m during each epoch are computed by accumulating the values τ_m and τ_m^2 during the epoch, computing the mean $\bar{\tau}$ and standard deviation s_τ according to the standard formulae

$$\bar{\tau} = \frac{1}{M} \sum \tau_m \qquad (8.4)$$

$$s_\tau^2 = \frac{\sum \tau_m^2 - \frac{(\sum \tau_m)^2}{M}}{M - 1} \qquad (8.5)$$

[cf. Quinn 1974 p.58] at the end of the epoch, and writing them out to the processed-data file.

A frequency histogram of the TIBB values τ_m is also compiled during the epoch and written to the processed-data file at the end of the epoch. The histogram consists of a set of numbers corresponding to "classes" of TIBB value, e.g. 0 - 0.5 seconds, 0.5 - 1.5 seconds,

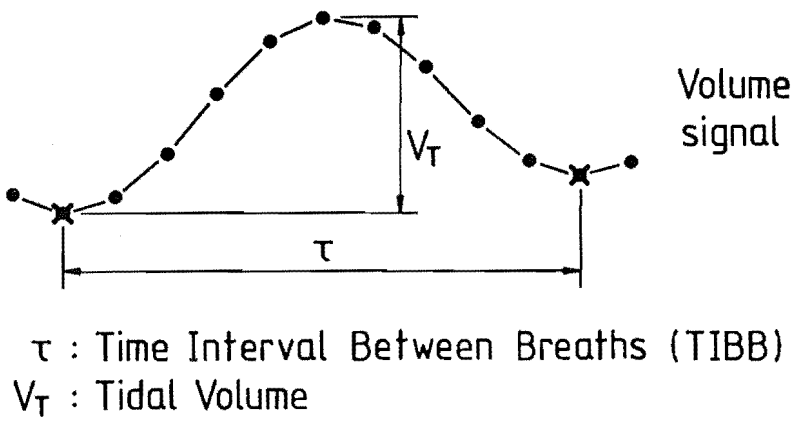


Figure 8.20: Determination of TIBB and tidal volume from the sampled signal.

1.5 - 2.5 seconds, etc. At the beginning of the epoch all of the numbers in the histogram are set to zero. During the epoch, as each breath is detected and its TIBB is computed, the number of the class to which the TIBB value τ_m belongs is incremented by one. At the end of the epoch each number in the histogram gives the frequency of occurrence of TIBB values of the corresponding class during the epoch. The frequency histogram provides an explicit estimate of the probability distribution of breath lengths during the epoch.

8.6.2.5 Tidal Volume

The tidal volume V_T is the difference in value between the peak and the trough of a breath (Figure 8.20). At the beginning of a breath, which occurs when a trough is detected in the sample stream, the software stores the value of the total volume signal. This value is subtracted from the value of the total volume signal when the following peak is detected, giving the tidal volume of the breath. The average tidal volume

$$\bar{V}_T = \frac{1}{M} \sum_{m}^{M} V_{Tm} \qquad (8.6)$$

of all breaths in the epoch is computed and written to the processed-data file at the end of the epoch.

8.6.2.6 Minute Ventilation

The average minute ventilation $\bar{\dot{V}}_E$ during the epoch is computed at the end of the epoch from

$$\bar{\dot{V}}_E = \frac{M}{T} \bar{V}_T \qquad (8.7)$$

and written to the processed-data file.

8.6.2.7 Chest/Abdomen Phase Index

The chest/abdomen phase index is a number expressing the

degree of synchrony of the chest and abdomen motions [cf. Sackner et al 1984b]. In normal breathing the chest and abdomen move in and out together, i.e. in phase. In some abnormal conditions such as obstructed or paradoxical breathing the chest and abdomen no longer

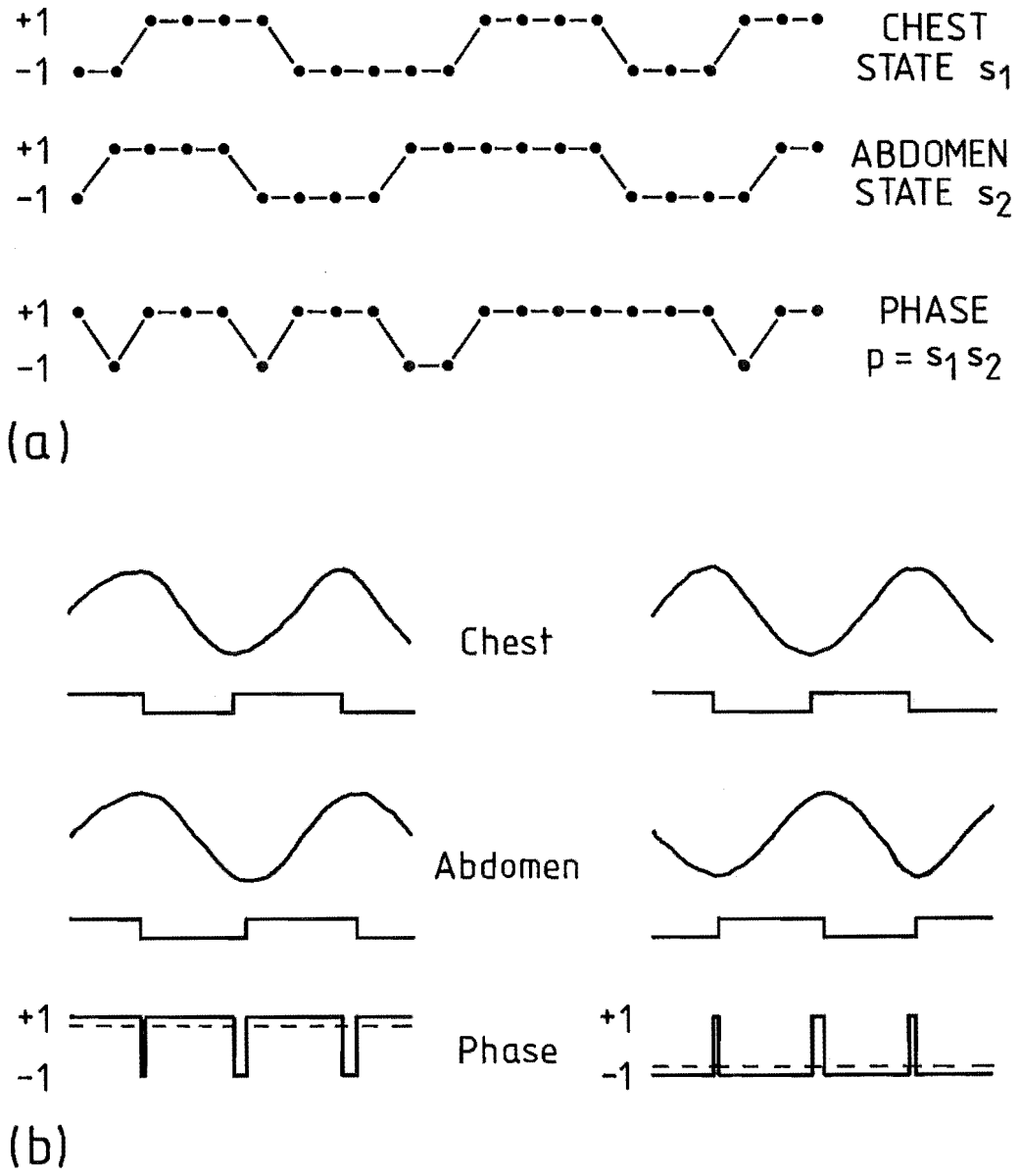


Figure 8.21: Computation and interpretation of the chest/abdomen phase index.

(a) Computing the phase $\{p_n\}$ from the states $\{s_{1n}\}$, $\{s_{2n}\}$.

(b) The phase index $\bar{p}$ (dashed line) is close to +1 for normal breathing (left) and close to -1 for paradoxical breathing (right).

move in phase. The purpose of the phase index is to identify the occurrence of such abnormal breathing motions, which can occur naturally from obstructive respiratory disease [Tobin et al 1983a] and have been associated with impaired blood oxygenation in opioid analgesia for postoperative pain relief (ø8.2.2).

For each sample the states s_{cn} and s_{an} of the chest and abdomen are compared and a phase p_n for the sample is defined to be +1 if the states are the same and -1 if they are different (see Figure 8.21a). If s_{cn} and s_{an} take the values ±1 (8.2) then

$$p_n = s_{cn}\, s_{an} \,. \tag{8.8}$$

The chest/abdomen phase index for the epoch is the average

$$\bar{p} = \frac{1}{N} \sum_{n}^{N} p_n \tag{8.9}$$

which is computed and written to the processed-data file at the end of the epoch. The value of $\bar{p}$ can vary between +1 when the chest and abdomen motions are perfectly in phase to -1 when the motions are perfectly out of phase (Figure 8.21b). Intermediate phasing of the motions, or a mixture of the extremes during an epoch, produces intermediate values of $\bar{p}$ (ø8.7).

8.6.3 DISPLAY

The processed-data files generated by the data analysis software contain information that can be directly displayed in graphical form. This is done using X/Y pen plotters at the Department of Electrical and Electronic Engineering, University of Canterbury. Two types of display are produced: a trend plot [cf. Theorell et al 1974, Wilson and Franks 1982] of the averaged values of the parameters of Section 8.6.2 for each epoch against time, and a stackplot of the frequency histograms of the TIBB in each epoch. These displays are intended to present the long-term trends in respiratory performance visually. When the clinical study of respiratory effects of analgesia (ø8.2.2) begins, it will be necessary to perform further data reduction on the processed-data files or the raw data to obtain

objective numerical descriptors of respiratory performance which can be statistically correlated against the type of analgesia administered.

8.6.3.1 Trend Plot

The "trend plot" is a graph of the values of each parameter's value vertically against the time of day horizontally (see Figure 8.22). The graphs of the several parameters are drawn one beneath another with the time scales aligned. From top downwards in Figure 8.22 these are TIBB, Minute Ventilation, Tidal Volume, and Phase Index. The horizontal bars in the TIBB graph show the mean TIBB in each epoch and the vertical bars denote ±1 standard deviation.

Trends in respiratory patterns with time are easily seen in the trend plot since information from a long interval is compactly presented and correlated changes in the different parameters are simultaneously visible. The minute ventilation and tidal volume are on arbitrary scales as the impedance plethysmograph is not yet quantitatively calibrated for volume. Therefore their values may be used in comparisons within a single recording but not necessarily between different recordings.

Below the graphs of breathing parameters, a number of graphs may be plotted which allow the quality of the recorded data to be assessed. These are the graphs B1, B2, F1, F2, and Noise Index in Figure 8.22. The B1, B2, F1, and F2 graphs show the mean value (horizontal bar) and range of values (vertical bar) of the Basal and Filtered signals from chest and abdomen in each epoch. Faulty electrical connections show as changes in the B1 and B2 graphs, which are normally constant at about half-scale. Noisy data are indicated by large ranges in the F1 and F2 graphs. The Noise Index graph shows the fraction of samples classified as artefacts (∅8.6.2.3) within each epoch and not used in computing the breathing parameters. A high noise index suggests that the remaining data, and hence the breathing parameters computed from them, may be unreliable.

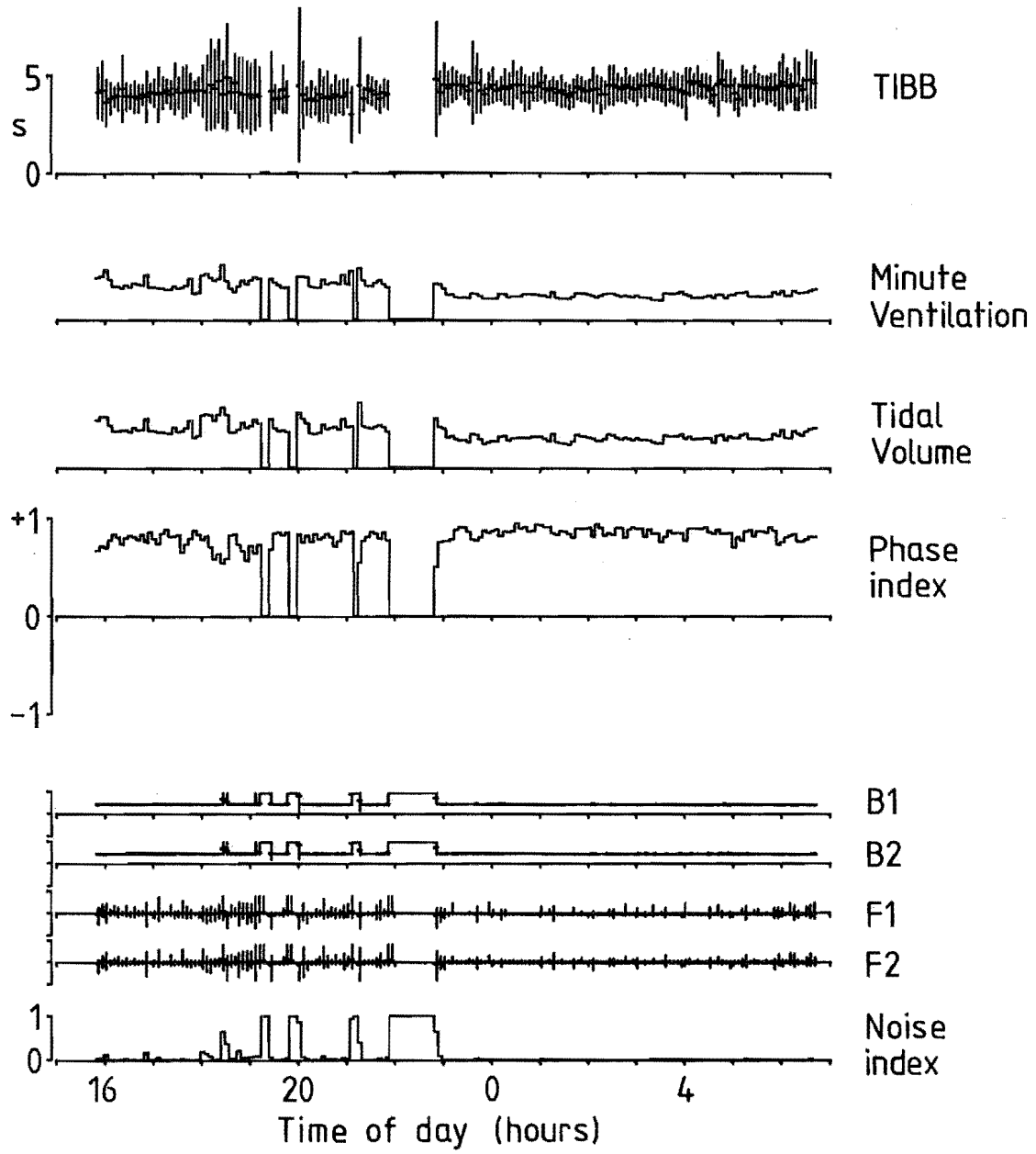


Figure 8.22: Trend plot of a clinical recording analysed over five minute epochs. The recording is of a middle-aged male patient before surgery.

8.6.3.2 TIBB Stackplot

Stacking the frequency histograms of TIBB values (ϕ8.6.2.4)

computed for each epoch one behind the other in a three-dimensional manner and drawing a perspective view of the stack in two dimensions generates a "stackplot" of the TIBB information, as illustrated in Figure 8.23 which was generated from the recording excerpted in Figure 8.16. The left-right horizontal coordinate is TIBB value, the vertical coordinate is frequency of occurrence of each TIBB value, and the front-back horizontal coordinate is time of day. Each data point's vertical coordinate is multiplied by the point's TIBB value (left-right horizontal coordinate) to obtain its vertical position in the stackplot. This weighting of the vertical coordinate compensates for the inverse relationship between the period of a breath and the

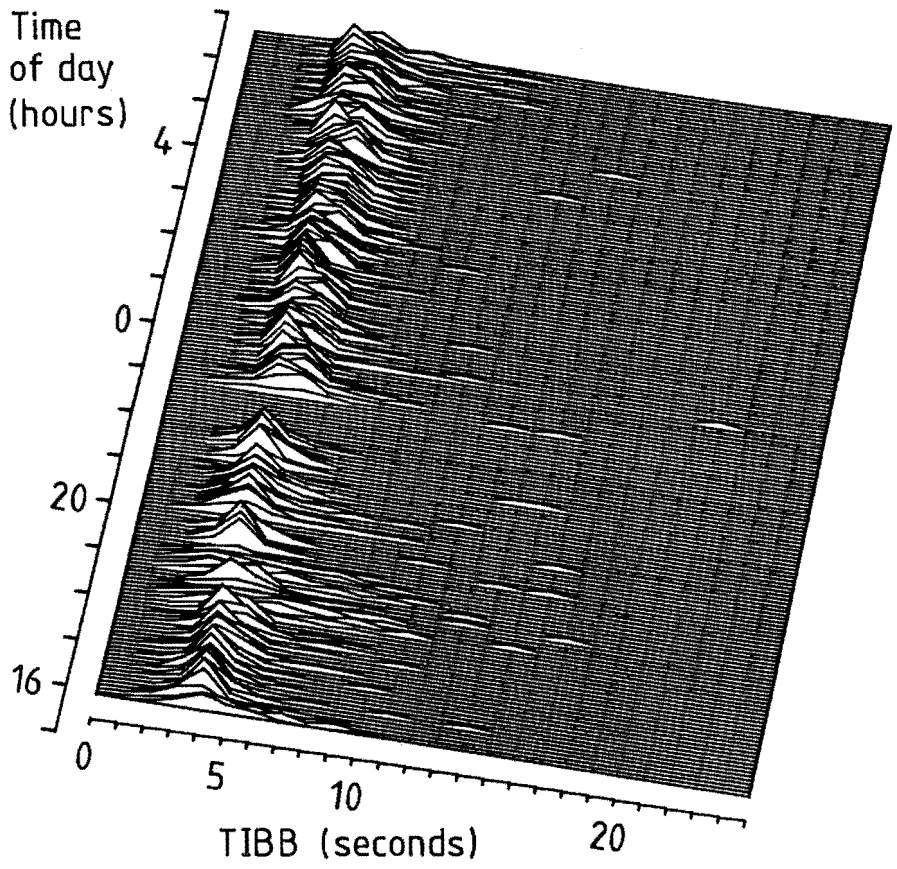


Figure 8.23: TIBB stackplot from the same data as Figure 8.22. Each slice in the plot is a frequency histogram of TIBB values in one epoch (see text). The left-right horizontal axis is TIBB value, the vertical axis is frequency of occurrence, and the front-back horizontal axis is time of day.

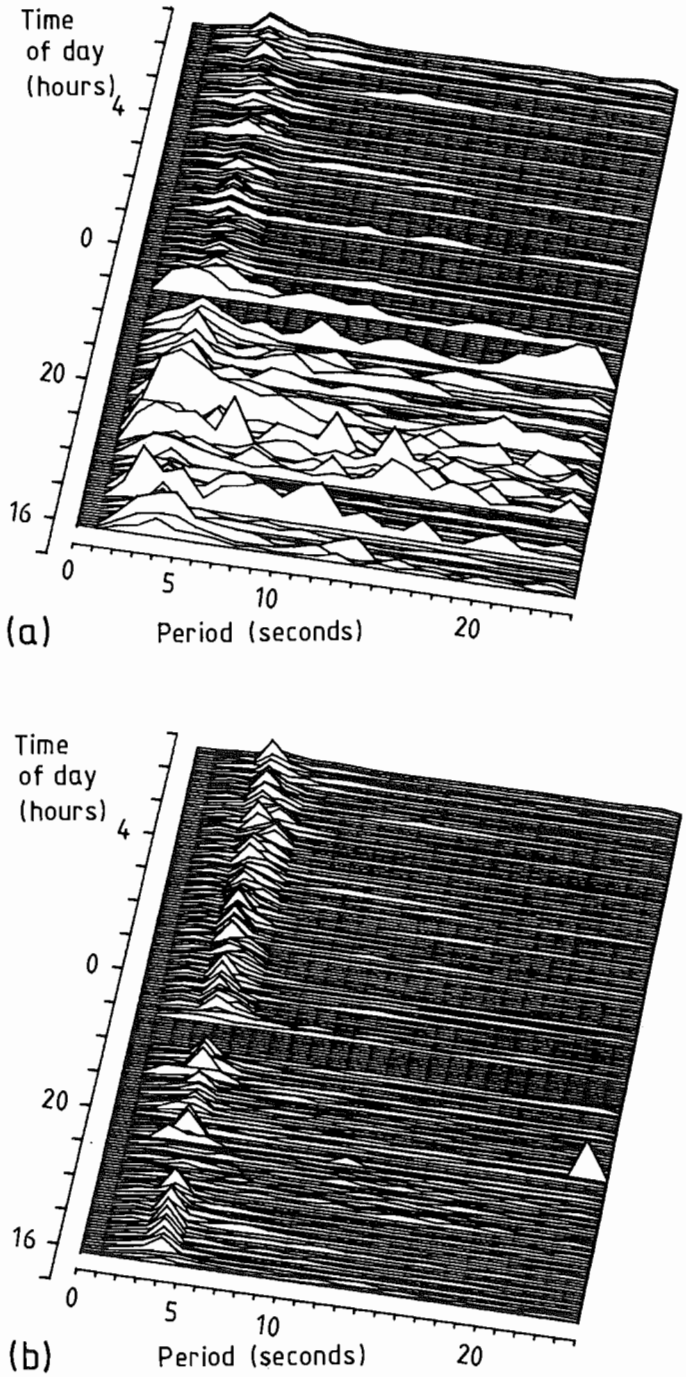


Figure 8.24: Stackplots of Fourier spectra from the same data as Figure 8.23, displayed on the same time scales as Figure 8.23. The vertical axis is spectral energy.

(a) Unnormalised: each slice is the spectral energy for one epoch.

(b) Normalised: each slice is normalised to give the slices constant area. Note the similarity to Figure 8.23.

number of such breaths which fit into an epoch of given length, so that steady breathing at any rate produces a peak of constant height in the display.

The stackplot display is visually striking and presents much information in a compact form. The mode (or most common value) of the TIBB distribution is the top of the ridge running the length of the stackplot, and variations in the mode with time are easily seen (ϕ8.7). The amount of spread in the TIBB distribution affects the height and width of the ridge. Breathing at a nearly constant rate produces a tall, narrow ridge whereas unsteady breathing or a noisy signal produces a broader, lower ridge.

The TIBB stackplot presents a primitive form of spectral analysis, with the frequency information on a linear period scale rather than the linear frequency scale of the conventional spectrogram [Altes 1980]. To demonstrate the relationship between the TIBB stackplot and conventional spectral analysis, Figure 8.24 shows two spectrograms, of the same recordings as in Figure 8.23, in which the power spectrum of each epoch has been computed using the Fast Fourier Transform [Bergland 1969] and the frequency coordinate has been transformed to a linear period scale like that of Figure 8.23. In Figure 8.24a the high amplitude of the signal in the noisy epochs obscures the quiet breathing in the display. In Figure 8.24b the area of each plot in the stack has been normalised, so that amplitude information in the signal is largely suppressed. Figure 8.24b is quite similar to the TIBB stackplot of Figure 8.23. The point of this comparison is that, if temporal information alone is required from the signal rather than amplitude information as well, the TIBB histograms produce results comparable to a Fourier spectral analysis [cf. Tucker 1967 ϕ3.4] and are easier to compute in real-time when the samples arrive at fixed, small intervals.

8.7 PRELIMINARY CLINICAL RESULTS

The respiratory monitoring system has been used in the wards of the Christchurch Hospital over the past year and a half to gather data and experience from which the system's hardware, operating

procedures, and data processing algorithms have been developed. Data have been recorded so far from approximately 25 patients. The amount and quality of the data are insufficient for conclusive statements about the respiratory effects of opioid analgesia, but preliminary observations can be made about the trends in respiratory behaviour that seem to be emerging.

Figures 8.25 and 8.26 show trend plots generated from preoperative and postoperative recordings, respectively, of a female patient of 73 years age undergoing a gastrectomy. Postoperative pain relief was provided by intermittent intramuscular injection of the opioid drug pethidine. The TIBB stackplots from the same recordings are shown in Figures 8.27 (preoperative) and 8.28 (postoperative). Sampling was at one second intervals. This pair of recordings shows clearly the changes between preoperative and postoperative breathing patterns which seem to be typical of the clinical data gathered so far.

Comparing the TIBB trend plots, Figures 8.25a and 8.26a, the preoperative breathing is at a relatively constant rate (apart from the noisy initial and corrupted final hours of the recording where the TIBB measurement is unreliable) whereas the postoperative breathing shows considerably greater long-term variations in TIBB. The preoperative and postoperative TIBB stackplots, Figures 8.27 and 8.28, show this strikingly. The increase in average TIBB in the postoperative state, which means slower breathing, is not typical of all the data gathered so far but the increased variability in the TIBB is. The origin of the short dips in the postoperative TIBB, marked * in Figure 8.26a, is unknown.

The minute ventilation and tidal volume, Figures 8.25b, 8.25c and 8.26b, 8.26c, seem to be less on average postoperatively than preoperatively. Unfortunately, since the inductance plethysmograph is not quantitatively calibrated for volume, this change may have come from altered placement of the transducer belts or an altered distribution of breathing motions between chest and abdomen coupled with incorrect scaling factors (ϕ7.4.3) rather than an actual decrease in the patient's ventilation.

The phase index shows a clear difference between the preoperative and postoperative states. In the preoperative recording (Figure 8.25d) the phase index (over the portion of the recording when the electrical connections to chest and abdomen were both intact) is comparatively steady and close to +1, indicating normal in-phase breathing motions almost all of the time. The postoperative phase index (Figure 8.26d) is much more variable and decreases to near or below zero on many occasions. This shows that the surgery or the opioid analgesia, or both, have significantly altered the patient's breathing in a potentially deleterious way. The phase index is not affected by the arbitrary calibration for volume of the inductance plethysmograph because it depends only on the relative times of inhalation and exhalation of the recorded signals (ϕ8.6.2.7).

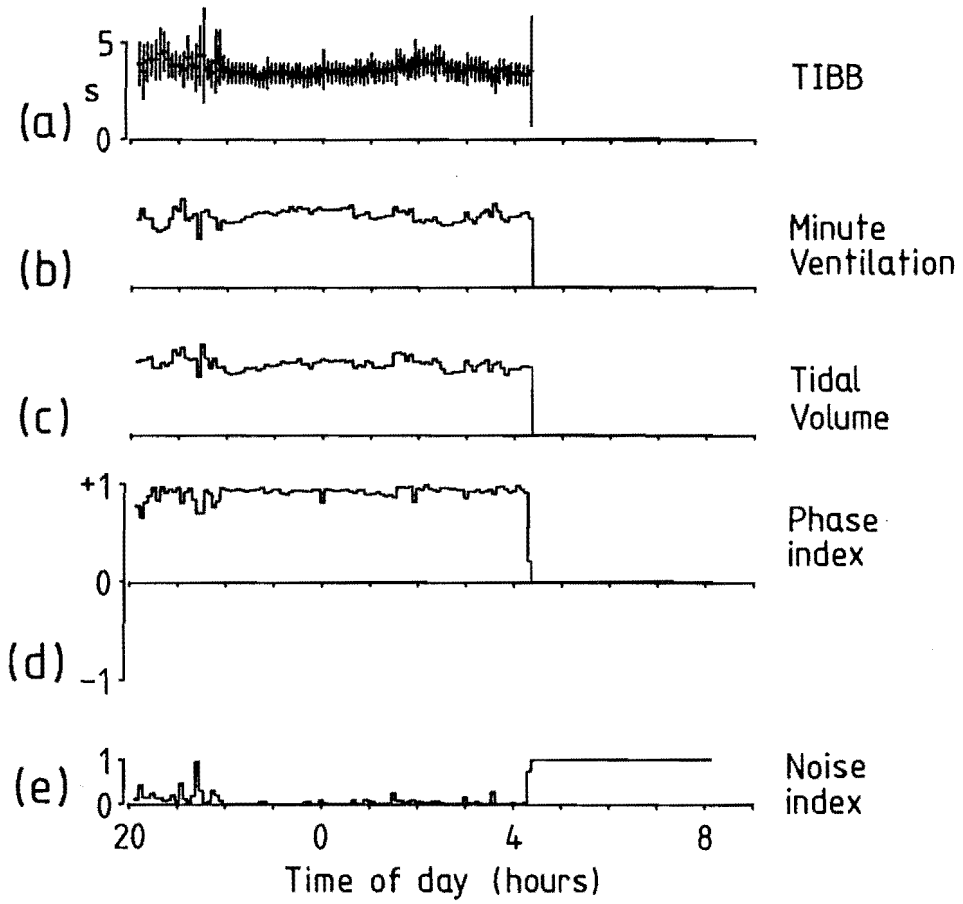


Figure 8.25: Trend plot of preoperative recording of a patient undergoing gastrectomy (see text). The abdomen belt came disconnected after 4pm.

An alternative technique of pain relief for some abdominal surgery is the intercostal nerve block [Dripps et al 1982 ϕ21], which affects only a small part of the nervous system and therefore may be expected to have less pronounced respiratory effects than general opioid analgesia. Figure 8.29 shows the postoperative phase indices of all patients recorded so far who were given intercostal blocks for postoperative pain relief. In each of these patients the postoperative phase index is comparatively uniform and close to +1. This suggests that it is the general opioid analgesia rather than the surgery which caused the changes in chest/abdomen phase in the patient of Figure 8.26, and that intercostal nerve blocks do not cause the same changes. The postoperative TIBB plots for these patients were unremarkable and hence are not shown here.

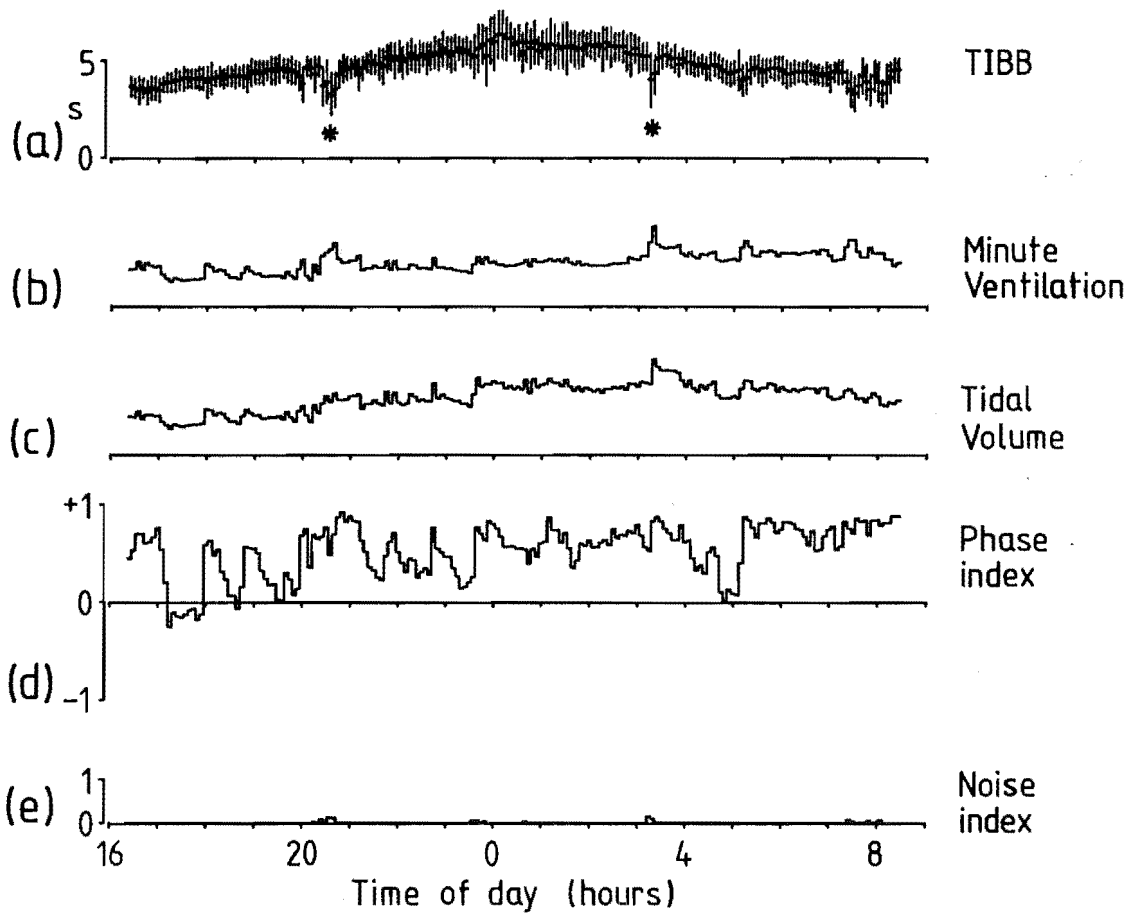


Figure 8.26: Trend plot of postoperative recording of a patient undergoing gastrectomy (see text). Postoperative pain relief was by intramuscular injection of an opioid pain-killer. * marks unexplained short dips in TIBB.

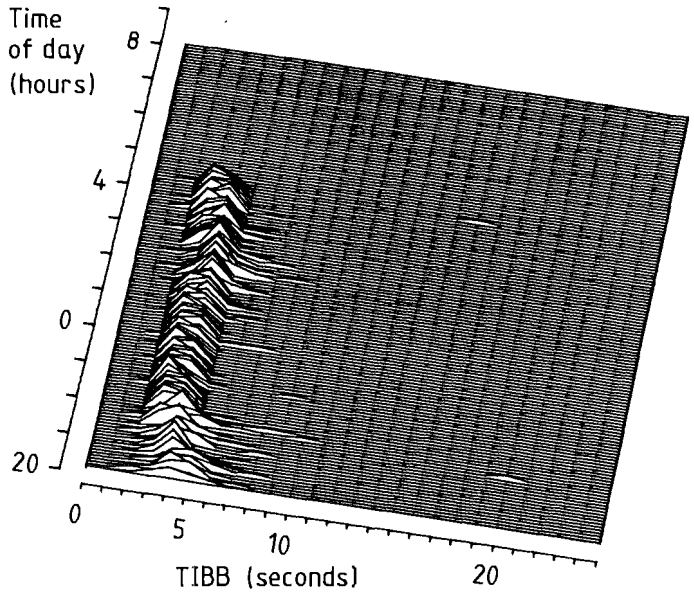


Figure 8.27: TIBB stackplot of preoperative recording (same data as Figure 8.25).

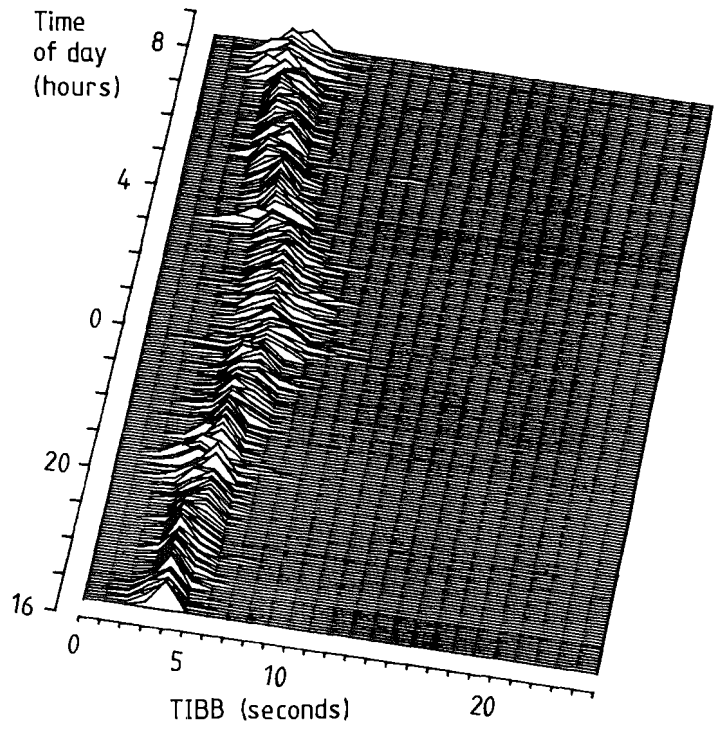


Figure 8.28: TIBB stackplot of postoperative recording (same data as Figure 8.26).

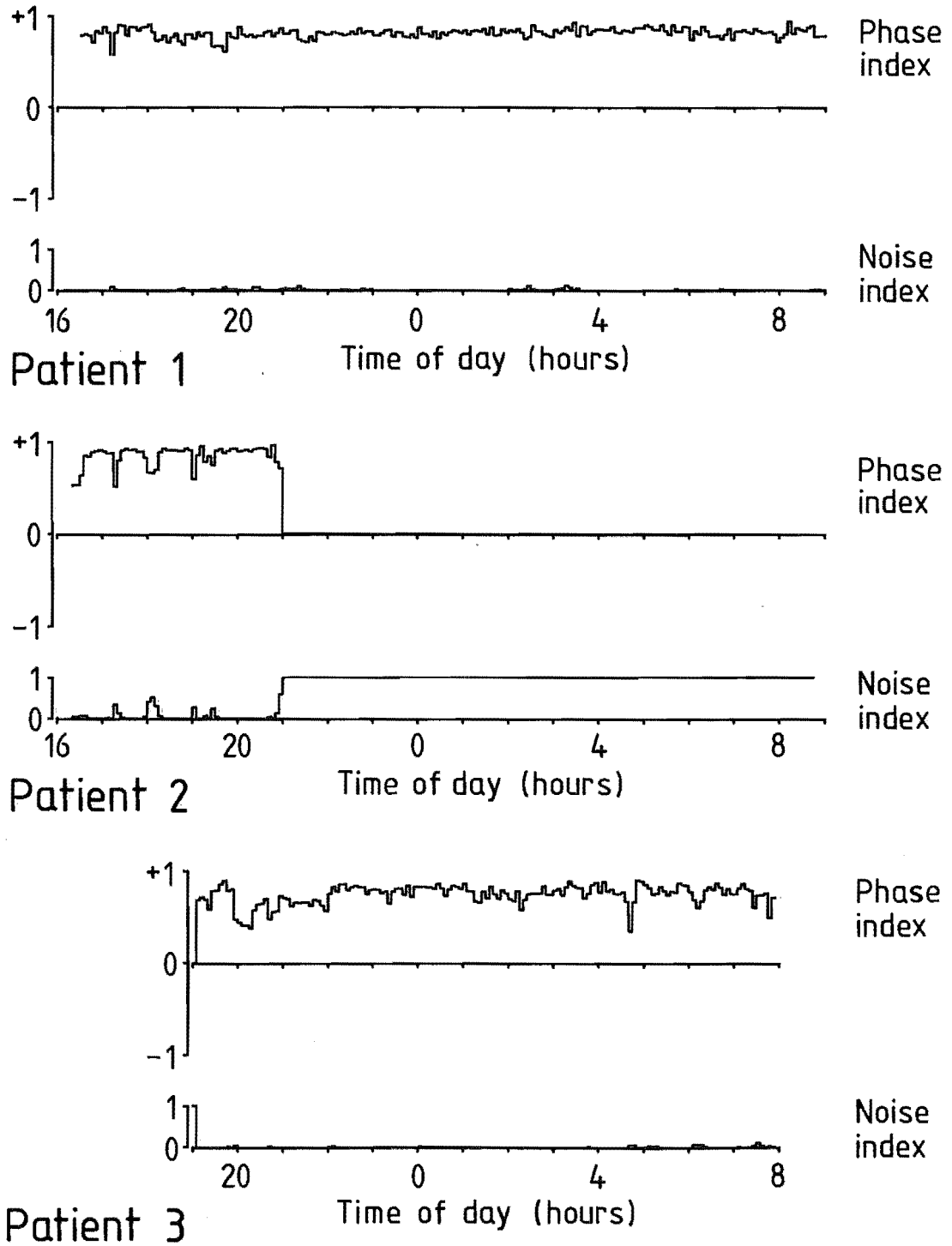


Figure 8.29: Phase and noise indices from postoperative recordings of three patients undergoing abdominal surgery and given intercostal nerve blocks for postoperative pain relief.

8.8 DISCUSSION

The inductance plethysmograph described in Section 8.4 has proven to be a good transducer for long-term clinical monitoring of respiration. It is comparatively comfortable to wear, does not interfere with the breathing of the patient, and delivers high quality recordings of chest and abdomen breathing motions except in the presence of gross body movement.

An inductance plethysmograph is easy to build, as the transducer belts require no special skills or elaborate equipment for their manufacture other than a sewing machine and a soldering iron, and the electronics need only measure an inductance or inductive impedance. Any electrical impedance plethysmograph should be able to function as an inductance plethysmograph, merely by connecting it to an elastic coil with an appropriate number of turns instead of directly to electrodes attached to the patient's body. This assumes that the electrical design of the impedance plethysmograph allows it to correctly measure an impedance which is predominantly inductive.

The transducer belts described in Section 8.4.1 have the important feature of being broken rather than continuous loops, allowing them to be easily fitted around the torso of patients after surgery. The extra mechanical complication incurred by the break does not importantly affect the reliability of the belts in use provided the electrical plug and socket connections are taped together so that they cannot easily be pulled free. The least reliable component of the belts has been the ribbon cable joining pieces (Figure 8.3a): the conductors inside the ribbon cable easily fatigue and break at the soldered joints to the 13-way flat plug (Figure 8.3c). To prevent this it has been necessary to encase the soldered joints in a rigid aluminium shroud filled with silicone rubber.

The relationship between the cross-sectional area of the transducer coils and their self-inductance (ϕ8.4.3.1) is smoothly curved, which means that over the small increment of area corresponding to the vital capacity range (Figure 8.7) the relationship is approximately linear. The relationship between lung volume and self-inductance has not been directly tested for the

plethysmograph described here, but by the same argument it can be expected to be approximately linear. The ±2.4% linearity found for cross-sectional area over the vital capacity range seems consistent on this basis with the 2% linearity for volume reported by Milledge and Stott [1977]. The measurements of change in coil self-inductance with shape reported in Section 8.4.3.2 support the assertion made by others [e.g. Watson 1979] that the relation between self-inductance and area can be treated as independent of coil shape, over the limited variations of shape that cross-sections of the human torso experience. The dependence of coil self-inductance on its detailed three-dimensional geometry (cf. the discontinuities in the fitted curves of Figures 8.7 and 8.9) is of no practical importance provided the belts are securely fixed to the patient during monitoring sessions.

To maintain accurate calibration of the inductance plethysmograph through an extended period of monitoring, it is obviously desirable for the positions of the belts on any patient to remain fixed. The mutual coupling between the coils (ø8.4.3.4) and the three-dimensional geometry of the coils (end of previous paragraph) are significant in this respect. Fixing the belts to the patient's skin with adhesive tape holds them securely but is unpleasant for hairy patients because removing the tape is painful unless expertly done. This is one of the practical aspects of the inductance transducer that deserves more attention. The manufacturers of the commercially produced Respitrace inductance plethysmograph originally attached both coils to a single elastic garment worn over the torso, but they later adopted an arrangement similar to that described in Section 8.4.1, with a separate elastic band for each coil [Sackner et al 1980b].

Quantitative calibration for volume of the inductance plethysmograph described here has not yet been attempted. The usefulness of the instrument in the postoperative respiratory monitoring application would be enhanced considerably by such calibration because quantitative estimates of the ventilation of the lungs, expressed for example as the minute ventilation, are an important measure of the performance of the mechanical part of the respiratory process. As discussed in Section 7.4.4, accurate determination of the calibration coefficients for chest and abdomen

requires special breathing manoeuvres (such as the isovolume manoeuvre) or significant variations in the relative contributions of chest and abdomen to the total changes in lung volume, to improve the conditioning of the mathematical system being solved. These requirements are unlikely to be met with patients undergoing surgery because the patients may be in pain and postoperatively they may not be conscious. Their cooperation cannot be assumed as is necessary for special breathing manoeuvres, and it may not be permissible to deliberately alter breathing patterns by changing their posture. Calibration for volume in this application is therefore likely to be difficult.

The eight bit resolution of the A/D conversion in the interface between the inductance plethysmograph and the microcomputer (ϕ8.5.1) is a significant limitation on the dynamic range of the system. The gain of the Filtered channels of the plethysmograph has been set so that the quietest breathing motions encountered in clinical monitoring are at least several least significant bits in amplitude. At this gain, moderately deep breathing saturates the A/D conversion and the plethysmograph output buffers in the Filtered channels. This has not caused any difficulties in the clinical monitoring application as patients resting in bed do not frequently breathe deeply, but it does affect the usefulness of the system as is in other applications such as pulmonary function testing where deep or vigorous breathing occurs. Should it prove desirable to extend the dynamic range of the analogue interface, there are two obvious ways to do this. The first is to install a higher resolution A/D converter (i.e. one with more bits). The programming language used for the respiratory monitoring software (ϕ8.5.2) has 16-bit integer arithmetic, but 10 or 12 bits of A/D conversion seems adequate for any clinical application of the system. The second and probably better way to achieve a greater dynamic range is to use an electronic compander [e.g. Wilson and Brown 1982] between the plethysmograph and the A/D converter to compress the dynamic range of the signals before conversion, and a lookup table in software to expand them again after conversion. Storage of raw data could be done in the 8-bit compressed form to save storage space.

The respiratory monitoring system's data processing currently

produces visual displays of parameters of respiratory performance averaged over five-minute epochs. The preliminary clinical results so far obtained (ϕ8.7) show that significant changes in breathing pattern do occur postoperatively and that the respiratory monitoring system can detect these changes. The proper clinical study of respiratory effects of analgesia for which the system was built will require the correlation of respiratory performance against factors such as the type of analgesia which is administered. To do this it will be necessary to define and compute objective numerical

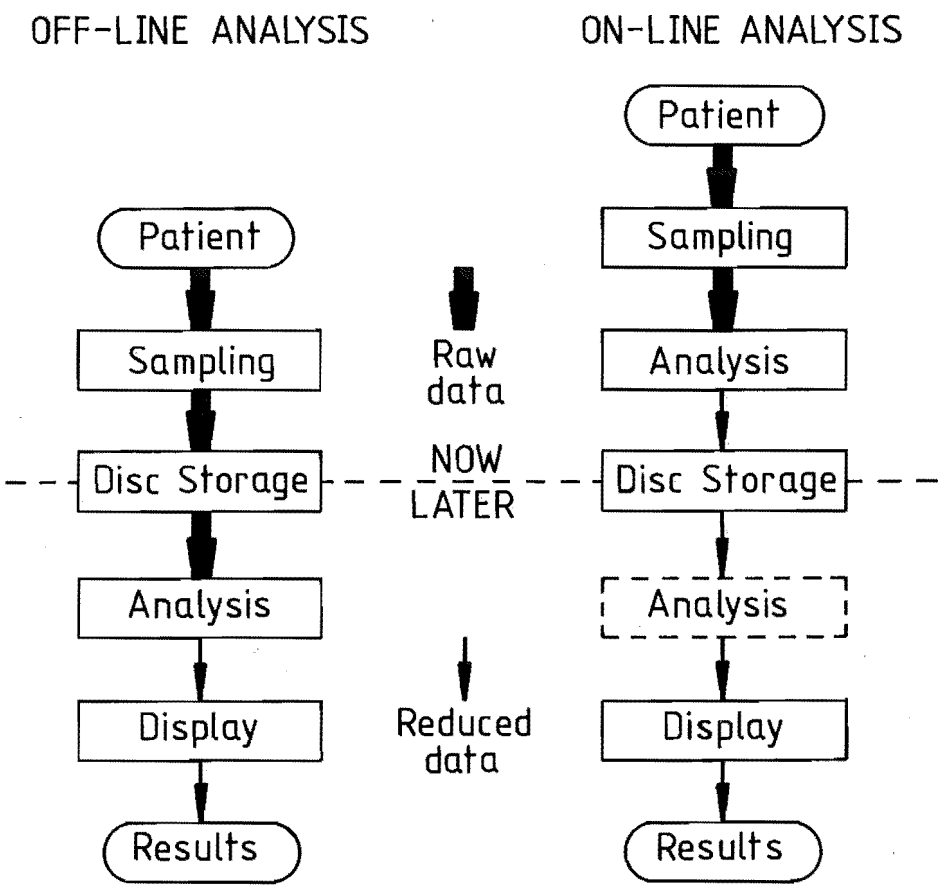


Figure 8.30: Data processing options for the respiratory monitoring system. The currently implemented data processing has the analysis done "off-line" (left), requiring the intermediate storage of a large volume of data. "On-line" processing (right) performs analysis in real time, which allows higher sample rates because the data is reduced before storage and allows "alarm" functions to be implemented.

descriptors of respiratory performance [cf. Wilson and Franks 1982, Sackner et al 1984b] which can be used as input to statistical tests. The statistical properties of these numerical descriptors will have to be carefully considered when applying statistical tests to them.

An important advance on the data processing currently implemented (ϕ8.6) is to do the data analysis in real-time, i.e. at the bedside as the data are acquired (see Figure 8.30). One benefit of this is the possible use of a high sample rate, since one does not have the restriction of storing all the samples on disc. More frequent sampling represents the signals with greater fidelity, which is likely to be necessary for more sophisticated tests for artefacts and processing of breath shape. The limited computational speed of the microcomputer imposes a limit on the sample rate and the complexity of the real-time data processing. However, the small bandwidth of respiratory waveforms [Hill 1973 p.2] allows more real-time data processing to be done by a given computer than in, say, ECG monitoring [Brydon 1976, cf. Bonneau et al 1983]. There are a number of situations to which an "alarm" monitor might be usefully applied, such as the sudden infant death syndrome [Golding et al 1985] and the adult sleep apnoea syndrome [Guilleminault et al 1976].

The respiratory monitoring system described in this chapter is potentially very versatile. Although designed and constructed for clinical monitoring of patients in studies of postoperative patient care, minor hardware enhancements and suitable software would make it a useful tool in other applications such as apnoea monitoring, pulmonary function testing, and exercise studies.

CHAPTER 9

AN AUTOMATED VENOUS OCCLUSION PLETHYSMOGRAPH

9.1 INTRODUCTION

This chapter describes a microprocessor-based automated venous occlusion plethysmograph, developed for studying blood circulation in the legs of patients undergoing hip replacement surgery. This instrument is the culmination of part of the Ph.D. research of a predecessor of mine [Seagar 1983a], and I completed its construction, programming, and commissioning.

Section 9.2 is a brief review of deep vein thrombosis (DVT) and of venous occlusion plethysmography, a non-invasive procedure which can detect DVT. Section 9.3 describes the hardware and operation of the automated venous occlusion plethysmography system. The data processing performed by the system is described in Section 9.4. Examples of results obtained with the system in clinical use are presented in Section 9.5. The usefulness of venous occlusion plethysmography and of the automated system are discussed in Section 9.6.

When I began work on the automated venous occlusion plethysmograph, its hardware had been specified and purchased and was almost completely assembled, and the data analysis algorithms described in Sections 9.4.3 to 9.4.7 already existed as part of a semi-manual system previously developed by Seagar. I completed the mechanical and electrical assembly of the automated system, producing the interfaces between the microcomputer and the printer, solenoid valve, and operator, finished the translation of the existing data analysis software from PDP FORTRAN into Intel PL/M-86, and designed and wrote the software for data acquisition, display of results, operator interface, and some other tasks within the system. My original contributions to the automated system are described in Sections 9.3.1, 9.3.2, 9.4.1, 9.4.2, and 9.4.8.

9.2 REVIEW

9.2.1 DEEP VEIN THROMBOSIS

Deep vein thrombosis (DVT) is an abnormal condition of the circulatory system in which blood clots or thrombi form in the large veins of the body, such as those of the lower legs. DVT is dangerous because the clots can become detached from their sites of formation and travel through the veins and the heart into the arteries leading to the lungs. A clot (or any foreign body) free in the bloodstream is called an embolus. The embolus eventually lodges in the arterial system of the lungs as the blood flows through vessels of decreasing size, producing the condition of pulmonary embolism and blocking the blood flow to part or all of the lungs. Depending on the size of the embolus and where it lodges in the pulmonary blood vessels, pulmonary embolism can be fatal to the patient, immediately or after a delay of up to several days [Guyton 1971 ϕ10].

Blood clotting is a mechanism to block leaks in the circulatory system arising for example from injury [Guyton 1971 ϕ10]. It is abnormal for blood to clot within an undamaged vessel as occurs with DVT. A set of three contributing factors to DVT has been termed Virchow's triad (see Figure 9.1), which forms a foundation for

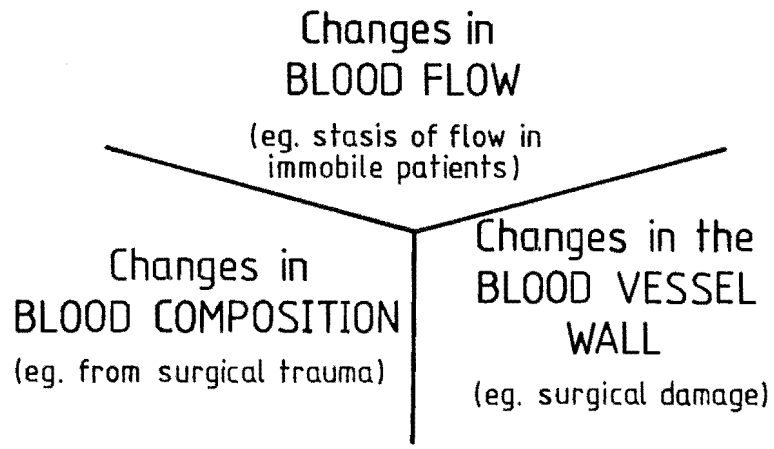


Figure 9.1: Virchow's Triad of contributing factors to deep vein thrombosis, with notes on possible effects of surgery.

modern descriptions of the causes of DVT [Calnan and Allenby 1975]. Deep vein thrombosis is a recognised risk to the patient undergoing surgery [--- 1978]. The risk is particularly great for hip surgery [Sevitt and Gallagher 1961]. Stasis of blood flow can occur, during surgery and postoperatively, because the patient is immobile. Furthermore, in hip replacement surgery the leg being operated on is dislocated and placed in an unnatural position. This can profoundly affect the blood circulation within the leg during the time the leg is dislocated, increasing the risk of DVT (Figure 9.1).

9.2.2 VENOUS OCCLUSION PLETHYSMOGRAPHY

Venous occlusion plethysmography (or VOP) is a non-invasive technique for measuring characteristics of the blood circulation within a part of the body [Sumner 1978a]. The veins leading out of the body part are occluded to prevent the outflow of blood from the body

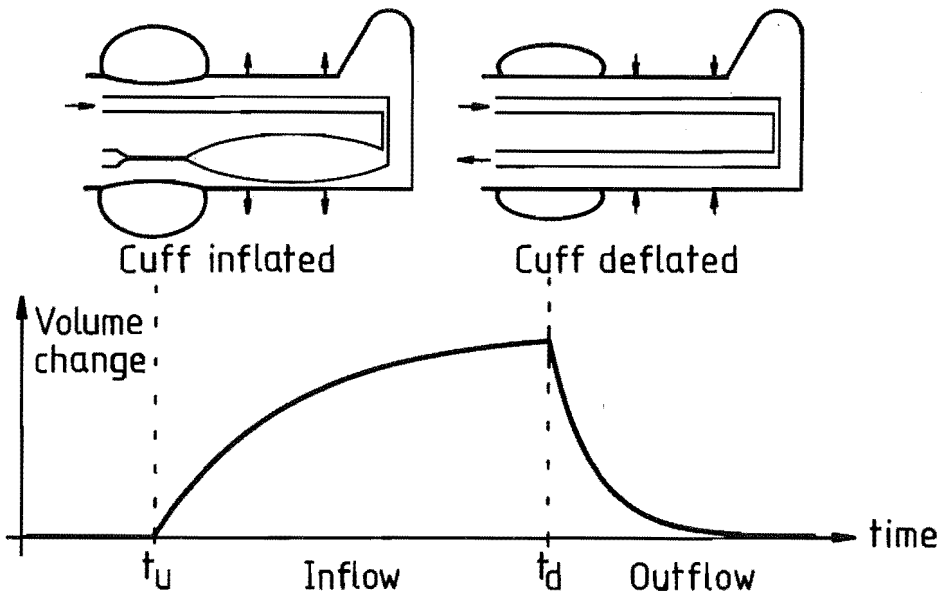


Figure 9.2: Volume changes in a limb during venous occlusion plethysmography (after Seagar [1983a]). When the cuff is inflated, blood collects in the veins and the limb's volume increases. When the cuff is deflated the excess blood flows out of the veins and the limb's volume decreases.

part and are then released again, causing pressure changes and consequent volume changes within the body part. On an arm or leg the occlusion is conveniently done by encircling the limb with a pneumatic cuff, inflating the cuff to a pressure sufficient to collapse the veins but not the arteries (in which the blood pressure is higher), and deflating the cuff again after a delay. The changes in volume of the limb during this process are measured (see Figure 9.2). The volume change with time is characteristic of the circulation within the limb and may be used to measure changes in the circulation, arising for example from deep vein thrombosis [Anderson and Wheeler 1979].

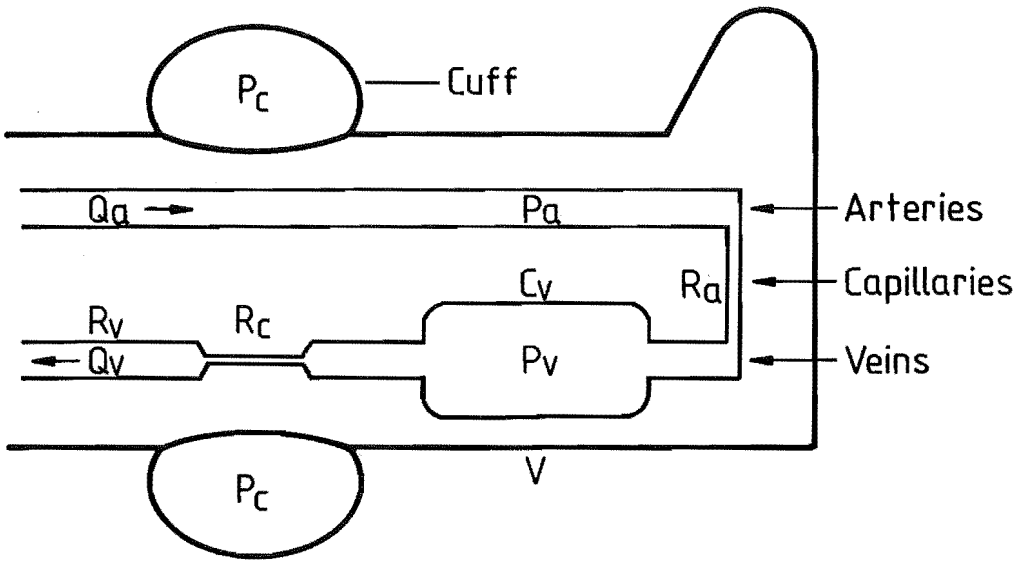
There are two commonly employed techniques for conveniently measuring changes in limb volume which circumvent the need to enclose the limb in an air- or water-displacement plethysmograph. The first of these is mercury strain gauge plethysmography [Sumner 1978b], which measures changes in the circumference of the limb. The second is electrical impedance plethysmography [Nyboer 1970, Anderson et al 1980], in which the electrical impedance of a section of the limb is measured and changes of this are related to changes in volume [Seagar 1983a ϕ1.3.2, Anderson et al 1980, Jaffrin and Vanhoutte 1979]. Impedance plethysmography can suffer from transient errors in determining limb volume change immediately after cuff deflation, due to changes in blood resistivity with flow [Wheeler 1978] (cf. ϕ9.4.4). However, comparative studies have shown that impedance plethysmography allows reasonably accurate determinations of limb blood flow during VOP and is in general more accurate than mercury strain-gauge plethysmography (see Mohapatra and Arenson [1979] and the references cited therein).

9.2.3 A SIMPLE MODEL OF LIMB CIRCULATION

Seagar et al [1984] describe a simple model of limb circulation which is used to interpret the changes of limb volume during venous occlusion plethysmography (see Figure 9.3). Since the model is an important part of the data analysis and presentation performed by the automated venous occlusion plethysmograph (ϕ9.4), it is described in detail here, following Seagar et al [1984] but with

slightly different notation.

The limb model represents all the blood vessels within the limb as a single vessel which is part artery, part capillaries (arterioles are lumped in with the capillaries), and part vein. The capillary and venous parts of the model vessel exhibit significant resistance to blood flow, and the venous part of the model vessel is



Q_a = arterial blood flow = $(P_a - P_v)/R_a$
Q_v = venous blood flow = $P_v/(R_c + R_v)$
P_a = arterial blood pressure
P_v = venous blood pressure
P_c = cuff pressure (= 0 when cuff deflated)
R_a = arteriolar resistance
R_v = venous resistance
R_c = cuff resistance (= 0 when cuff deflated)
C_v = venous compliance
V = limb volume

$$\frac{dV}{dt} = Q_a - Q_v = C_v \frac{dP_v}{dt}$$

Figure 9.3: Model of limb circulation during venous occlusion plethysmography (after Seagar et al [1984]).

compliant, i.e. changes in blood pressure are accompanied by changes of volume of the vein. The volume change in the entire model limb is equal to the volume change in the vein. Inflation of a pneumatic cuff encircling the limb for venous occlusion plethysmography (∅9.2.2) adds a large cuff resistance R_c to the venous resistance R_v, causing the pressure within the vein to rise so that the vein expands.

Blood flow through the vessel occurs at pressures P and flows Q, which are different at different parts of the vessel (Figure 9.3). The resistances of the vessel to flow are independent of flow so that the pressure drop ΔP across a resistance R is related to flow Q by

$$\Delta P = QR \ . \qquad (9.1)$$

Similarly, venous compliance C_v is independent of pressure within the vein so that volume change ΔV is related to pressure change ΔP in the vein by

$$\Delta V = C_v \Delta P_v \ . \qquad (9.2)$$

The physiological quantities of the limb model are compactly stated as follows:

SYMBOL	NAME	UNITS OF MEASUREMENT (for example)
V	Volume	cm^3
Q	Flow	cm^3/s
P	Pressure	mmHg
R	Resistance	$mmHg/(cm^3/s)$
C	Compliance	$cm^3/mmHg$.

(9.3)

An electrical equivalent circuit of the model is shown in Figure 9.4, with the correspondence between the electrical circuit parameters and the physiological model parameters of Figure 9.3 shown.

The response of the model to inflation and deflation of the cuff is, in terms of changes of volume of the limb (cf. changes in charge on the capacitor in Figure 9.4), of exponential form after the cuff is inflated and of different exponential form after the cuff is

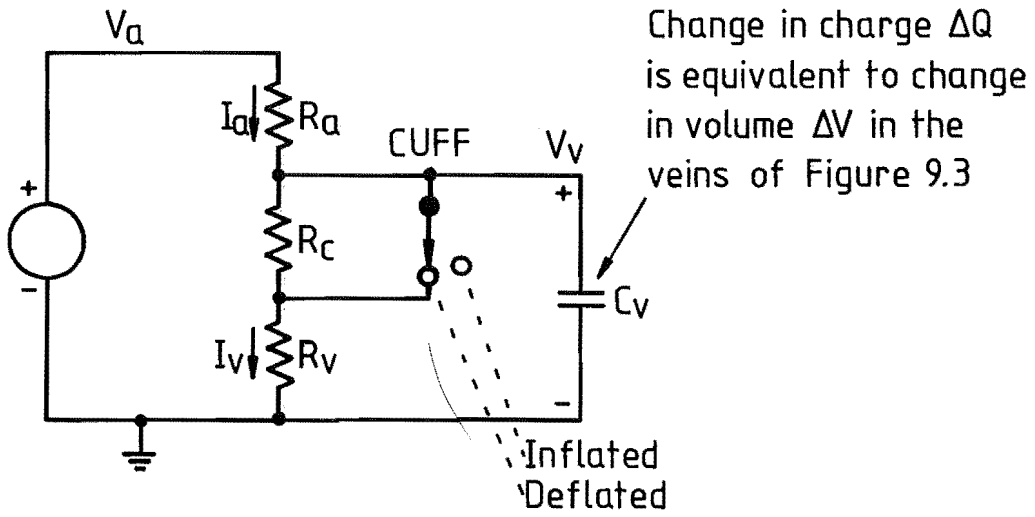
deflated again (cf. Figure 9.2).

The volume changes in venous occlusion plethysmography are commonly expressed as fractions of the average total volume [Sumner 1978a], in units of "%" or "$cm^3/100cm^3$" say. It is then convenient to define "specific" quantities

$$C_u = C_v/V \ , \tag{9.4a}$$

$$R' = R.V \tag{9.4b}$$

[Seagar et al 1984]. Using the specific quantities, (9.3) becomes



Physiological	Electrical
Q Flow	I Current
P Pressure	V Voltage
V Volume = $\int Q.dt$	Q Charge = $\int I.dt$
C Compliance = $\Delta V/\Delta P$	C Capacitance = $\Delta Q/\Delta V$
R Resistance	R Resistance

Figure 9.4: Electrical equivalent circuit for the limb model of Figure 9.3, and correspondence of the electrical quantities with the physiological model quantities.

SYMBOL	NAME	UNITS OF MEASUREMENT (for example)
$\Delta V/V$	Relative volume change	% or $cm^3/100cm^3$
Q/V	Specific flow	$(cm^3/100cm^3)/s$
P	Pressure	mmHg
R'	Specific resistance	$mmHg/((cm^3/100cm^3)/s)$
C_u	Specific venous compliance	$(cm^3/100cm^3)/mmHg$. (9.5)

The relative volume change of the model limb during venous occlusion plethysmography is given by

$$\frac{\Delta V}{V}(t) = \begin{cases} 0 & \text{for } t \leq t_i\,, \quad (9.6a) \\ C_u\,\Delta P_m \left(1 - e^{\frac{-(t-t_i)}{C_u R_1'}} \right) & \text{for } t_i \leq t \leq t_o\,, \quad (9.6b) \\ V_t\, e^{\frac{-(t-t_o)}{C_u R_2'}} & \text{for } t \geq t_o \quad (9.6c) \end{cases}$$

where t_i and t_o are the times at which inflow and outflow begin, which are ideally equal to the instants t_u and t_d of cuff inflation and deflation respectively (Figure 9.2), and

$$R_1' = R_a' \| (R_c' + R_v') = \frac{R_a'(R_c' + R_v')}{R_a' + R_c' + R_v'} \quad \text{(cuff inflated)}, \quad (9.7a)$$

$$R_2' = R_a' \| R_v' = \frac{R_a' R_v'}{R_a' + R_v'} \quad \text{(cuff deflated)}, \quad (9.7b)$$

$$\Delta P_m = P_v\Big|_{t\to\infty} - P_v\Big|_{t=t_i} = P_m - P_i = P_a \frac{R_1'}{R_a'} - P_a \frac{R_2'}{R_a'}\,, \quad (9.7c)$$

$$V_t = \frac{\Delta V}{V}(t)\Big|_{t=t_o} \quad (9.7d)$$

(cf. Seagar et al [1984], in which $\Delta P_m \approx P_m$ has been used in equation (3)).

The model is invoked by fitting exponential functions to the

inflow and outflow portions of the measured volume change function (Figure 9.2) and calculating the model parameters (e.g. Q_a/V, C_u, R_v') from the parameters of the fitted exponential curves (ϕ9.4.7). The model parameters are descriptive of the limb circulation, and have the advantage over other arbitrary measures of the volume change function [Seagar et al 1984] of being more closely related to the actual physiology of the limb.

9.3 THE AUTOMATED VENOUS OCCLUSION PLETHYSMOGRAPH

9.3.1 HARDWARE

The hardware of the automated venous occlusion plethysmograph (also referred to here as "the system") comprises a pneumatic circuit and cuffs, an electrical impedance plethysmograph, a chart recorder, and a Processor Box and printer (see Figure 9.5). The chart recorder

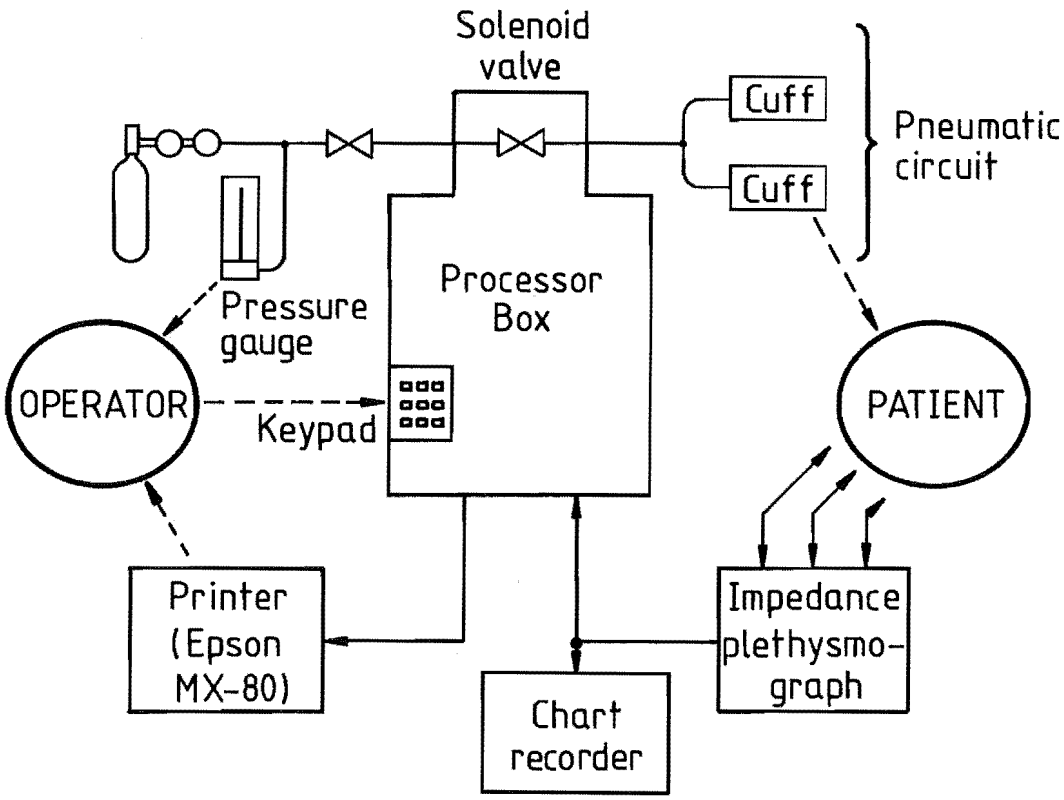


Figure 9.5: Block diagram of the automated venous occlusion plethysmograph.

is not necessary to the operation of the automated system but has been retained for continuity with the earlier semi-manual system (ϕ9.1). The automated VOP is designed to make simultaneous measurements on both legs of a patient during hip replacement surgery, so there are two pneumatic cuffs and two channels in the impedance plethysmograph, one for each leg. The processor box contains a microcomputer and controls the inflation and deflation of the cuffs, samples the electrical outputs of the impedance plethysmograph, computes and prints out the limb model parameters and other results of the measurements, and executes self-test routines, under commands entered by the system's operator through a keypad mounted on the processor box.

The microcomputer within the processor box (see Figure 9.6) consists of an Intel iSBC 88/25 processor board [Intel 1981] containing an 8088 microprocessor and 8087 numeric coprocessor, a 64K byte dynamic RAM board [Sinton 1979], and a 12-bit A/D converter board [Analog Devices 1978], connected by the Intel Multibus [Barthmaier 1979, Boberg 1980]. The processor box also contains power supplies and interfaces between the microcomputer and the impedance plethysmograph, solenoid valve, keypad, and printer.

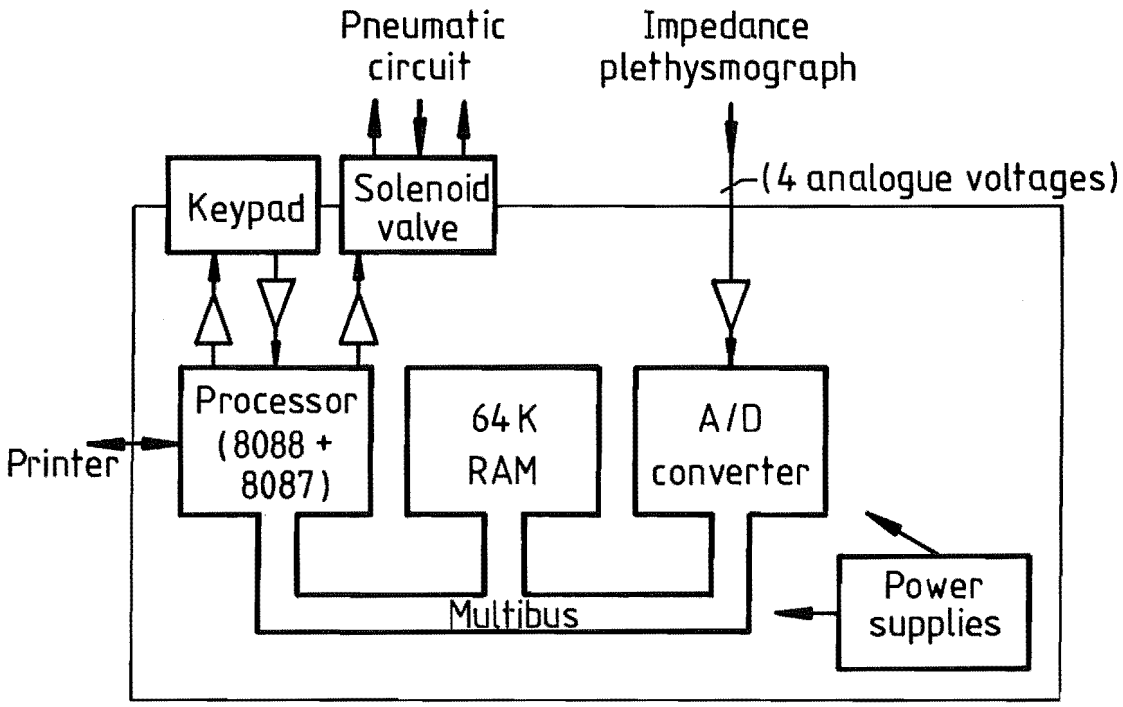


Figure 9.6: Block diagram of the automated VOP processor box.

The software for the system is written in PL/M-86 [Intel 1978], comprising approximately 1600 PL/M-86 statements and compiling to approximately 18K bytes of executable code which resides in EPROM on the iSBC 88/25.

9.3.2 OPERATING CONTROLS

The operating controls of the automated venous occlusion plethysmograph consist, apart from mains switches on the electrical equipment and various manual valves in the pneumatic circuit, of a keypad of sixteen keys and a large red Stop button on the processor box (see Figure 9.7). The Stop button applies a hardware reset to the microcomputer. It is not normally used. The keypad has twelve numeric keys and four command keys which cause the system to perform specific actions when they are pressed.

MEA (Measure) causes the automated VOP to make one set of measurements (ϕ9.4.2). Samples representing the volume changes in the legs are taken at regular intervals and stored in the system's memory. The pneumatic cuffs are inflated and deflated at appropriate times. There is room in memory for two sets of measurements, i.e. two successive uses of the Measure command.

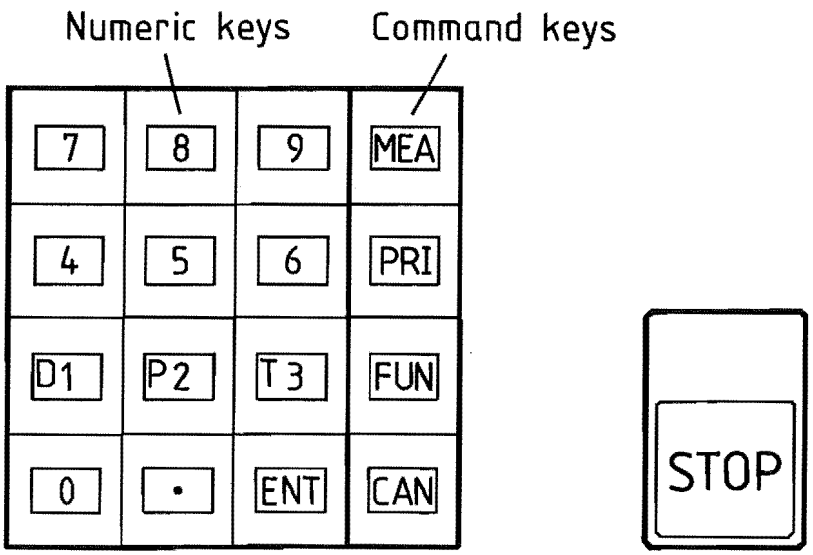


Figure 9.7: Operating controls of the automated VOP processor box.

PRI (Print) causes the automated VOP to compute and print out (ϕϕ 9.4.3 - 9.4.8) the results from the least recent measurement stored in memory.

FUN (Function) used in combination with the numeric keys 1 (Discard), 2 (Parameters alter), or 3 (Test), allows (1) a measurement to be discarded without processing, (2) the parameter values used in the processing to be altered, or (3) self-testing routines to be executed which exercise the hardware of the system.

CAN (Cancel) aborts the execution of the Measure or Print commands, quickly returning the system to readiness for a new command.

The numeric keys, together with the Cancel key for cancelling bad entries, are used with the Function 2 (parameters alter) command to enter numbers into the system.

9.3.3 CLINICAL USE OF THE AUTOMATED VOP

In clinical use of the automated VOP to monitor leg blood flow during hip replacement surgery, measurements (a single measurement records inflow and outflow curves, as depicted in Figure 9.2, for each leg) are taken before surgery, at intervals during surgery, and on occasions during some weeks following surgery. The procedure for the measurements at each of these times is essentially similar.

The impedance plethysmograph electrodes are of self-adhesive aluminium tape applied around the circumference of the patient's legs with electrolyte jelly. A constant current source connected to an electrode at each ankle drives 2mA p-p AC through both legs. The AC voltages appearing across pairs of electrodes around each calf are measured. The pneumatic cuffs are placed just above the knees. For measurements made during surgery, elastic stockings which reduce the risk of deep vein thrombosis by diverting blood from surface veins to deep veins in the leg [Calnan and Allenby 1975] are fitted over the electrodes. This arrangement remains in place during surgery.

A measurement is made by pressing the Measure key (ϕϕ 9.3.2,

9.4.2). The processed results of the measurement are obtained by pressing the Print key (ϕϕ 9.3.2, 9.4.3 - 9.4.8).

One measurement takes approximately 80 seconds (if the default inflow and outflow times are used), during which time surgery is halted to keep the patient's legs still and thereby avoid motion artefacts. The processing and printing of results takes considerably longer, about 5 minutes, due almost entirely to the slow speed of the printer (processing takes about 5 seconds and printing about $2\frac{1}{2}$ minutes for each leg). The Measure and Print functions are separate commands and the processor box maintains two memory buffers so that two Measurements can be made followed by two Print commands. This allows closely spaced pairs of measurements to be made, for example to observe the effects of a particular action on the legs. Table 9.1 shows the times at which VOP measurements were initiated in a typical

Surgical Procedure	VOP Measurement Label	Start time	Time between measurements (minutes)
	Pre induction	11:10am	
Initial anaesthesia			28
	Pre incision	11:38am	
Hip cut open			15
	Pre dislocation	11:53am	
Joint dislocated			6
	Post dislocation	11:59am	
Acetabulum (cup) fitted			32
	Post acetabulum	12:31pm	
Femoral prosthesis fitted			3
	Post shaft	12:34pm	
Joint reassembled			2
	Post relocation	12:36pm	
Cut sewn shut			16
	End operation	12:52pm	

Table 9.1: Times of VOP measurement during a typical total hip replacement operation.

total hip replacement operation. The two-measurement capability reduces the delaying of normal surgical procedure during the closely spaced post acetabulum to post relocation measurements by allowing the processing and printing of a measurement to be held until after the post relocation measurement.

9.4 DATA PROCESSING

9.4.1 TURN-ON INITIALISATION

When the processor box is turned on, or reset using the Stop button, software is executed which initialises the hardware of the processor box, performs self-tests, and sets default parameter values for the data acquisition and data analysis described in the following subsections. During the initialisation procedure "INIT" is printed on the printer, and when initialisation is complete the processor box enters a "ready" state in which "READY> " is printed as a prompt and the processor waits for a command from the operator (∅9.3.2).

9.4.2 DATA ACQUISITION

Data acquisition is initiated by pressing the Measure button. Real-time interrupts generated on the iSBC 88/25 board (∅9.3.1) regularly activate an interrupt service routine (ISR). The ISR is a special subroutine activated by a hardware signal instead of a software Call instruction. The ISR periodically samples the impedance plethysmograph output voltages through the A/D converter and stores the sample values into RAM. The ISR also energises and releases the solenoid valve to inflate and deflate the pneumatic cuffs at appropriate times. At the end of the measurement period, the ISR halts the real-time interrupts and signals completion of measurement to the main program. Meanwhile, the main program executes a loop, writing the values of the most recent sample to the printer in a graphical format simulating the traces of a chart recording (see Figure 9.8). This output gives immediate visual feedback to the system's operator of the quality of the measurement. The loop

terminates when the completion of measurement signal from the ISR occurs, and the main program then returns to its ready state and waits for the next command from the operator. The default values of the sampling interval and the times relative to the start of measurement of cuff inflation, cuff deflation, and measurement termination are established in the turn-on initialisation as follows:

SAMPLE INTERVAL	1.0 seconds
CUFF INFLATION TIME	2 seconds
CUFF DEFLATION TIME	47 seconds
MEASUREMENT END TIME	77 seconds.

The default values give 45 seconds of inflow measurement and 30 seconds of outflow measurement (cf. Figure 9.2). These values may be altered by the operator using the Function 2 (parameters alter) command (ϕ9.3.2).

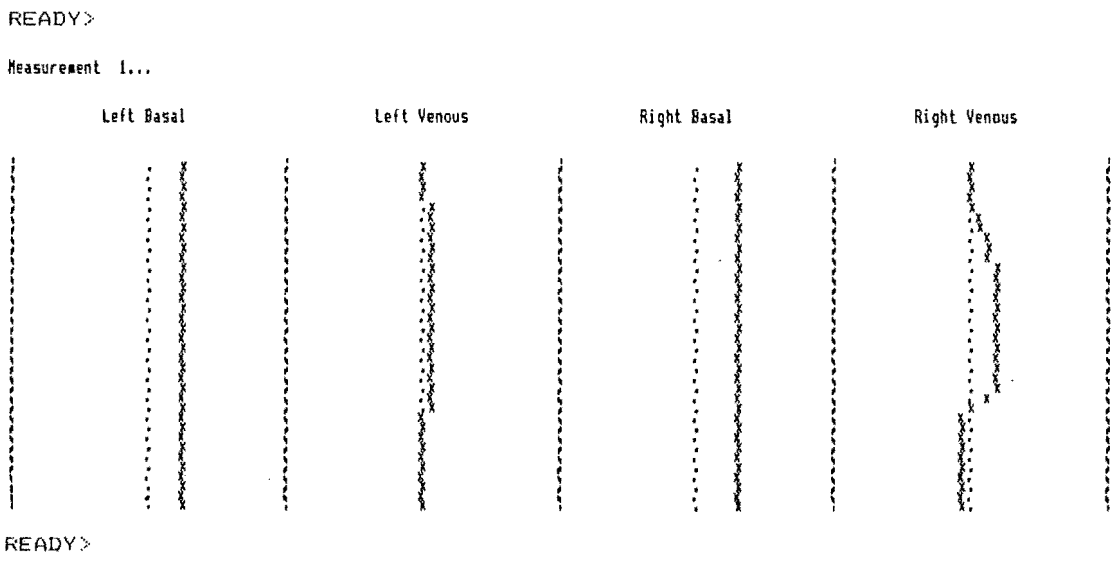


Figure 9.8: Real-time display of the impedance plethysmograph voltages during measurement. This display simulates the output of a four-channel chart recorder, with ±10V full scale horizontally for each channel and time running vertically down the page. Crosses mark the measured voltages, dots mark zero volts, and the broken vertical lines delimit the four channels.

The measurements are made and stored in terms of limb impedance Z(t) and impedance changes ΔZ(t) as measured by the impedance plethysmograph. These data are converted into relative volume changes during the analysis phase (∅9.4.3). Since the volume changes and consequent impedance changes in the limb are comparatively small (of the order of 1% of the average volume or impedance, cf. Figures 9.11, 9.12 in ∅9.5), the impedance plethysmograph provides two electrical outputs for each leg to allow both Z(t) and ΔZ(t) to be accurately represented within the dynamic range of the microcomputer's A/D converter and the chart recorder. The first ("basal") output voltage (cf. Figure 9.8) is directly proportional to Z(t). The second ("venous") output voltage, representing ΔZ(t), is obtained from the basal voltage via a high-pass RC network and amplifier providing high gain with a DC block. For each measurement, the venous voltages for each sample and the average basal voltage over all samples during the measurement period for each leg are scaled into units of ohms and stored.

9.4.3 SCALING TO RELATIVE VOLUME CHANGE

The processing and printing of results from a measurement, initiated by the Print command (∅9.3.2), begins by computing relative volume change ΔV/V(t) for each leg using the stored impedance data from the measurement phase (∅9.4.2). The stored average impedance from the basal signals gives Z, the average limb impedance during the measurement, directly. The impedance changes with time ΔZ(t) are obtained from the stored venous samples by deconvolving from them the impulse response of the high-pass RC networks in the impedance plethysmograph's venous channels [Seagar 1983a Appendix 2]. The relative volume changes ΔV/V(t) are then computed from the impedance data using the approximation

$$\frac{\Delta V}{V}(t) = \frac{-\rho_b}{\rho_t} \frac{\Delta Z(t)}{Z} \qquad (9.8)$$

[Seagar 1983a ∅1.3.2, Anderson et al 1980], where ρ_b and ρ_t are the blood and tissue resistivities respectively. Equation (9.8) is derived from a cylindrical model of a limb consisting of blood and

tissue, in which the blood volume but not the tissue volume changes and the length of the limb remains constant. Default values for ρ_b and ρ_t of 150Ωcm and 240Ωcm respectively are used. These may be changed by the operator using the Function 2 (parameters alter) command.

9.4.4 FINDING INFLOW AND OUTFLOW

The instants t_i and t_o at which inflow and outflow begin are ideally equal to the instants t_u and t_d of inflation and deflation of the cuffs (Figure 9.2). However, in practice the volume increase and decrease associated with inflow and outflow, respectively, are not

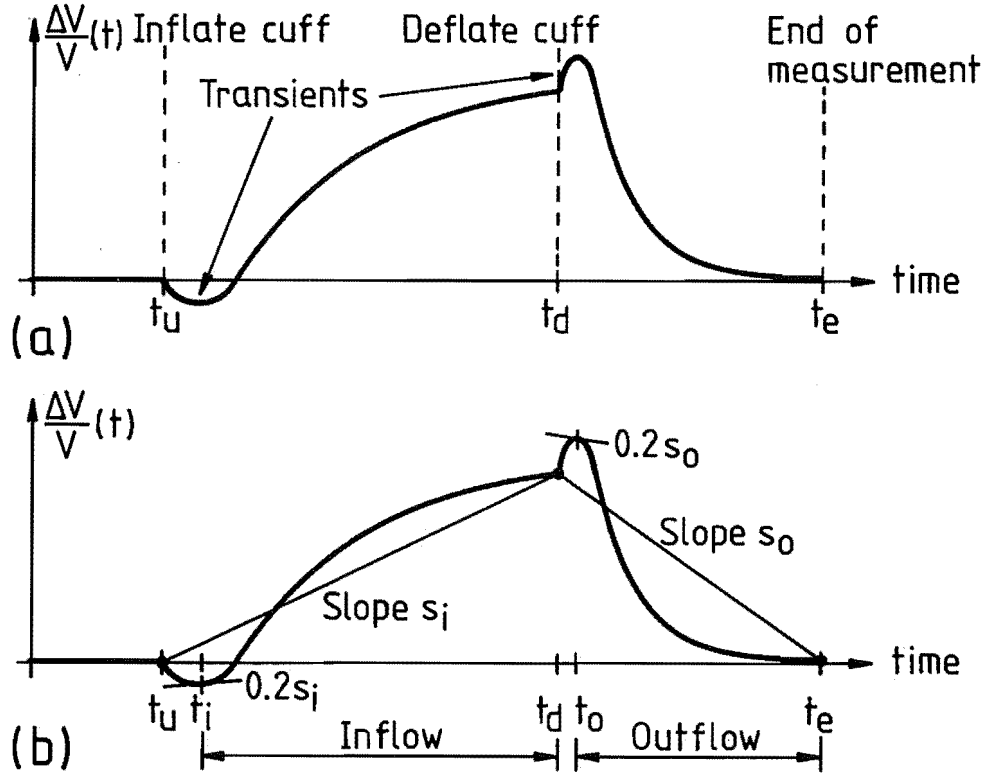


Figure 9.9: Finding the start of inflow and outflow.

(a) Transient changes in $\frac{\Delta V}{V}(t)$ occur immediately after inflation and deflation of the cuffs.

(b) Triggering to avoid the inflation and deflation transients. Inflow is defined as the interval $[t_i; t_d]$ and outflow as the interval $[t_o; t_e]$.

necessarily apparent immediately after the inflation or deflation of the cuffs [Seagar 1983b]. The measurements may show the behaviour immediately after cuff inflation and deflation depicted in exaggerated form in Figure 9.9a, due perhaps to tissue displacement near the cuffs or changes in blood resistivity with flow rate [Seagar 1983b]. When fitting exponential curves to the measurements, it is advantageous to do so over portions of the measurements that do not include the inflation and deflation transients.

The software incorporates a "trigger" mechanism to define the beginnings of inflow and outflow after the inflation and deflation transients. The trigger exploits the oppositeness of direction of change of the transients from the expected volume changes. Using inflow as an example, the average slope of the $\Delta V/V(t)$ curve over the interval $[t_u;t_d]$ is computed, and 20% of this slope is defined as a threshold. The point within the interval $[t_u;t_d]$ at which the slope of $\Delta V/V(t)$ first becomes greater than the threshold is recognised as the beginning of inflow (Figure 9.9b). For outflow the same process is performed over the interval $[t_d;t_e]$ (t_e is the time of end of measurement), and since the expected average slope is negative the beginning of outflow is recognised where the slope first becomes less than the threshold.

The volume changes of the limb being operated on in hip surgery are sometimes very small [Seagar 1983b] (cf. Figure 9.12 in ∅9.5), so that the beginning of inflow and outflow cannot be accurately identified from the volume change measurements of the operative limb. The volume changes of the other limb are usually normal, however (cf. Figure 9.12). The software trigger described here is applied to the sum of the volume changes of both limbs, and the instants t_i and t_o so found are used for both limbs. The intervals $[t_i;t_d]$ and $[t_o;t_e]$ are defined as the durations of inflow and outflow, respectively, and the exponential curves are fitted to the measurements (∅9.4.6) over these intervals.

9.4.5 REMOVING BASELINE DRIFT

The limbs should ideally return to the same volume after cuff

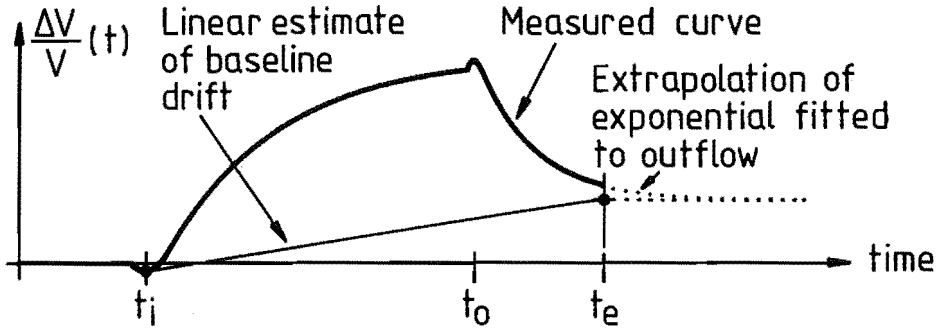


Figure 9.10: Correction of relative volume change for baseline drift.

deflation as they had before cuff inflation. In practice small deviations from this, of up to say 10% of the volume changes from the venous occlusion itself, may occur [Seagar 1983b] (see Figure 9.10). The software removes this baseline drift, assuming its rate to be constant over the measurement period, by fitting an exponential function to the outflow curve to determine the final resting volume, drawing a straight line from this volume at t_e to the volume at the beginning of inflow at t_i (Figure 9.10), and subtracting this linear estimate of the baseline drift from the measured $\Delta V/V(t)$ to get a corrected $\Delta V/V(t)$.

9.4.6 FITTING EXPONENTIALS TO INFLOW AND OUTFLOW

Exponential functions are fitted to the corrected samples (ϕ9.4.5) of $\Delta V/V(t)$ to obtain the parameters A_1, B_1, C_1 and A_2, B_2, C_2 in the equations

$$\text{INFLOW} \quad \frac{\Delta V}{V}(t) = A_1 + B_1 e^{-C_1(t-t_i)}, \quad t_i \leq t \leq t_d, \qquad (9.9)$$

$$\text{OUTFLOW} \quad \frac{\Delta V}{V}(t) = A_2 + B_2 e^{-C_2(t-t_o)}, \quad t_o \leq t \leq t_e. \qquad (9.10)$$

Each exponential is fitted by forming an initial estimate of the

parameters A_1, B_1, C_1 (inflow) or A_2, B_2, C_2 (outflow) and then iteratively refining the estimate to minimise the RMS difference between the exponential and the measured data over the interval $[t_i;t_d]$ (inflow) or $[t_o;t_e]$ (outflow) using a modified Newton algorithm [Seagar et al 1984]. The numerical algorithms used are described by Seagar [1983a].

9.4.7 INVOKING THE LIMB MODEL

The limb model of Section 9.2.3 is invoked by equating (9.9) and (9.10) with (9.6b) and (9.6c) and solving for the model parameters Q_a/V, C_u, and R_v' [Seagar et al 1984]. The maximum pressure rise ΔP_m (9.7) is required for these calculations. Experiments have shown [Seagar et al 1984, Seagar 1983b] that $\Delta P_m \approx P_m$ is approximately half of the cuff pressure P_c over the range of cuff pressures normally used in venous occlusion plethysmography. The software therefore uses the approximation

$$\Delta P_m = 0.5 P_c \ . \qquad (9.11)$$

The default value used for P_c is 60mmHg. If a different cuff pressure is used, P_c can be altered with the Function 2 (parameters alter) command.

The model parameters are calculated using

INITIAL SPECIFIC INFLOW $$\left.\frac{Q_a}{V}\right|_{t=t_i} = -B_1 C_1 \ , \qquad (9.12a)$$

SPECIFIC VENOUS COMPLIANCE $$C_u = \frac{A_1}{P_m} \ , \qquad (9.12b)$$

SPECIFIC VENOUS RESISTANCE $$R_v' = \frac{1}{C_u C_2} \ . \qquad (9.12c)$$

A number of other parameters of the volume change curve are also calculated, to allow direct comparison with the other measures that have been used in venous occlusion plethysmography [Seagar et al 1984]. These are as follows:

MAXIMUM SPECIFIC OUTFLOW $\left.\frac{Q_v}{V}\right|_{t=t_o} = B_1 C_2$, (9.13a)

OUTFLOW TIME CONSTANT $\tau_o = \frac{1}{C_2}$, (9.13b)

MAXIMUM VOLUME RISE $\frac{\Delta V}{V}(t_d)$ from the measured data, (9.13c)

RELATIVE RISE AFTER 2s OUTFLOW $\frac{\frac{\Delta V}{V}(t_d+2)}{\frac{\Delta V}{V}(t_d)}$ from the measured data, (9.13d)

RISE AFTER 3s OUTFLOW $\frac{\Delta V}{V}(t_d+3)$ from the measured data. (9.13e)

The maximum volume rise and rise after 3s outflow are displayed (ϕ9.4.8) in units of 0.04cc/100cc. They may be reexpressed as percentage volume changes (cc/100cc) by dividing their numerical values by 25.

9.4.8 DISPLAY

After the data analysis is complete, the processor box produces a page of print-out for each leg which presents the results of the analysis (examples of this print-out are given in Figure 9.11 in ϕ9.5). The print-out includes numerical values of measurement conditions, parameter values assumed in the analysis, and results from the analysis, and gives a graphical picture of the relative volume change $\Delta V/V(t)$ in percent units against time, after deconvolution (ϕ9.4.3) and removal of baseline drift (ϕ9.4.5). In the graph, measured data points appear as crosses "x" and the fitted curves appear as dots "·". Where these coincide a cross is printed. The RMS difference between the measured and fitted curves, in percent units, is displayed graphically as a vertical bar at the top right of the graph and numerically below the graph, giving a quantitative assessment of the quality of fit of the model response to the actual limb response.

This space intentionally left blank.

9.5 EXAMPLES OF CLINICAL RESULTS

This section presents examples of output from the automated venous occlusion plethysmograph in situations where it has provided clinically significant information.

The first example (see Figure 9.11) illustrates the effects of deep vein thrombosis on blood flow within the legs as interpreted by venous occlusion plethysmography. Figure 9.11a (left leg) is of a leg affected by proximal DVT (the thrombus is in the upper leg, "downstream" in the veins from the VOP measurement site), and Figure 9.11b (right leg) is of a normal leg. The blood flow and volume change from venous occlusion are smaller in the affected leg, as is clearly evident from the percentage volume change curves of Figure 9.11. Comparing the numerical values of the electrical model parameters for the two legs, the affected leg (Figure 9.11a) has reduced initial specific inflow, greatly reduced maximum specific outflow, and increased specific venous resistance. These changes are consistent with partial blockage of the veins carrying blood away from the left lower leg. Experience with venous occlusion plethysmography in Christchurch hospitals has shown that the value of the specific venous compliance (styled "capacitance" in Figure 9.11) is by itself not a reliable indicator of the presence or absence of DVT.

The second example (see Figure 9.12) shows changes of blood circulation during surgery in the legs of a patient undergoing total hip replacement surgery. For a time during surgery the limb being operated on (the right leg in Figure 9.12) is twisted to an unnatural position for the fitting of the femoral prosthesis. This deranges the blood vessels at the hip and can greatly affect blood flow within the limb. The top, middle, and bottom rows of Figure 9.12 show volume change curves from VOP measurements on both legs at the post acetabulum, post shaft, and post relocation stages of surgery (cf. Table 9.1 in ϕ9.3.3) respectively. The right leg (right hand curve in each row) was the operative limb. It was in a comparatively normal position for the top and bottom rows and in the unnatural position for the middle row of Figure 9.12.

The VOP measurements show that the blood flow within the right

Left Leg, measurement 1.

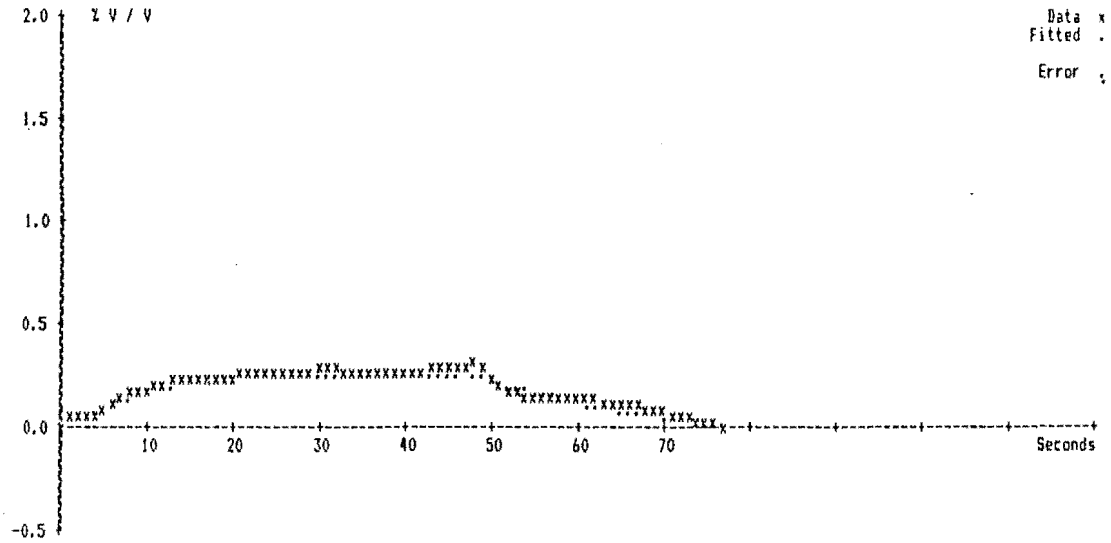
Measurement conditions --

```
Sampling interval :  1.0 seconds
Cuff inflated at  :    2 seconds
Cuff deflated at  :   47 seconds
Cuff pressure assumed      :   60 mm Hg
Blood resistivity assumed  :  150 ohm.cm
Tissue resistivity assumed :  240 ohm.cm
```

Measured results --

```
Inflow detected at       :    3 seconds
Outflow detected at      :   48 seconds
Average impedance of leg :   26 ohms
```

Percentage volume changes ...



RMS absolute error in fit : 0.016 %V/V

Electrical Model parameters --

```
Initial specific inflow     :   1.88  cc/100cc/minute
Maximum specific outflow    :    0.6  cc/100cc/minute
Specific Venous Capacitance :   0.93  cc/100cc/100mmHg
Venous "resistance"         : 27.088  100mmHg/(cc/100cc/sec)
Outflow timeconstant        :  25.10  sec
```

Additional parameters --

```
Maximum volume rise              :    7.6  0.04cc/100cc
Relative rise after 2s outflow   :  0.826  0.04cc/10cc
```

(a)

Figure 9.11: Output of the automated venous occlusion plethysmograph showing the effects of proximal deep vein thrombosis in a patient's left leg (a), compared to the right leg (b) which was unaffected by DVT.

Right Leg, measurement 1.

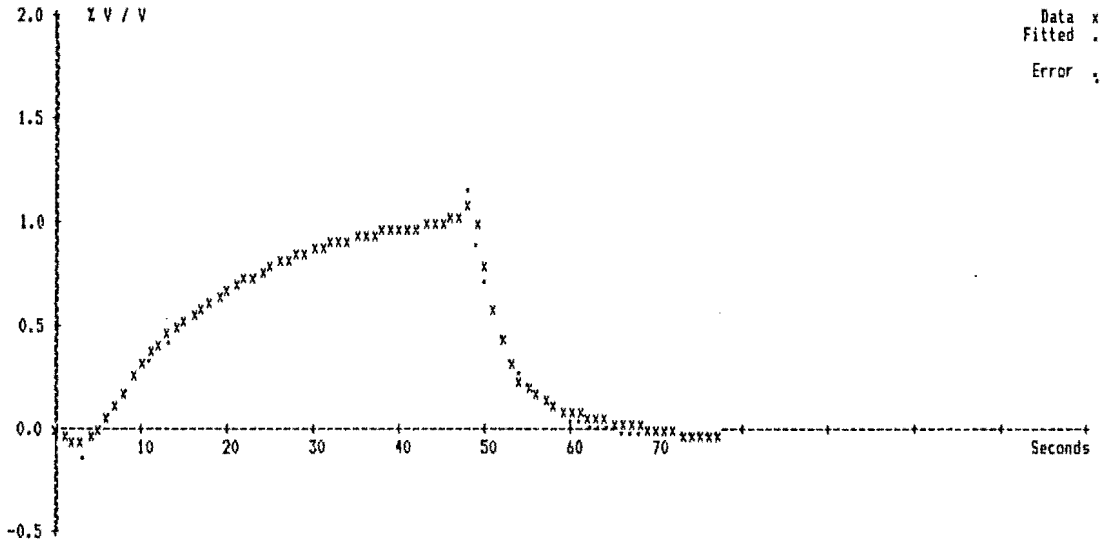
Measurement conditions --

```
Sampling interval :  1.0 seconds
Cuff inflated at  :    2 seconds
Cuff deflated at  :   47 seconds
Cuff pressure assumed      :   60 mm Hg
Blood resistivity assumed  :  150 ohm.cm
Tissue resistivity assumed :  240 ohm.cm
```

Measured results --

```
Inflow detected at       :    3 seconds
Outflow detected at      :   48 seconds
Average impedance of leg :   35 ohms
```

Percentage volume changes ...



RMS absolute error in fit : 0.022 %V/V

Electrical Model parameters --

```
Initial specific inflow     :   4.51  cc/100cc/minute
Maximum specific outflow    :   16.8  cc/100cc/minute
Specific Venous Capacitance :   3.67  cc/100cc/100mmHg
Venous "resistance"         :  1.178  100mmHg/(cc/100cc/sec)
Outflow timeconstant        :   4.32  sec
```

Additional parameters --

```
Maximum volume rise            :   26.1  0.04cc/100cc
Relative rise after 2s outflow :  0.959  cc/cc
Rise after 3s outflow          :   20.0  0.04cc/10cc
```

(b)

Figure 9.11 (continued)

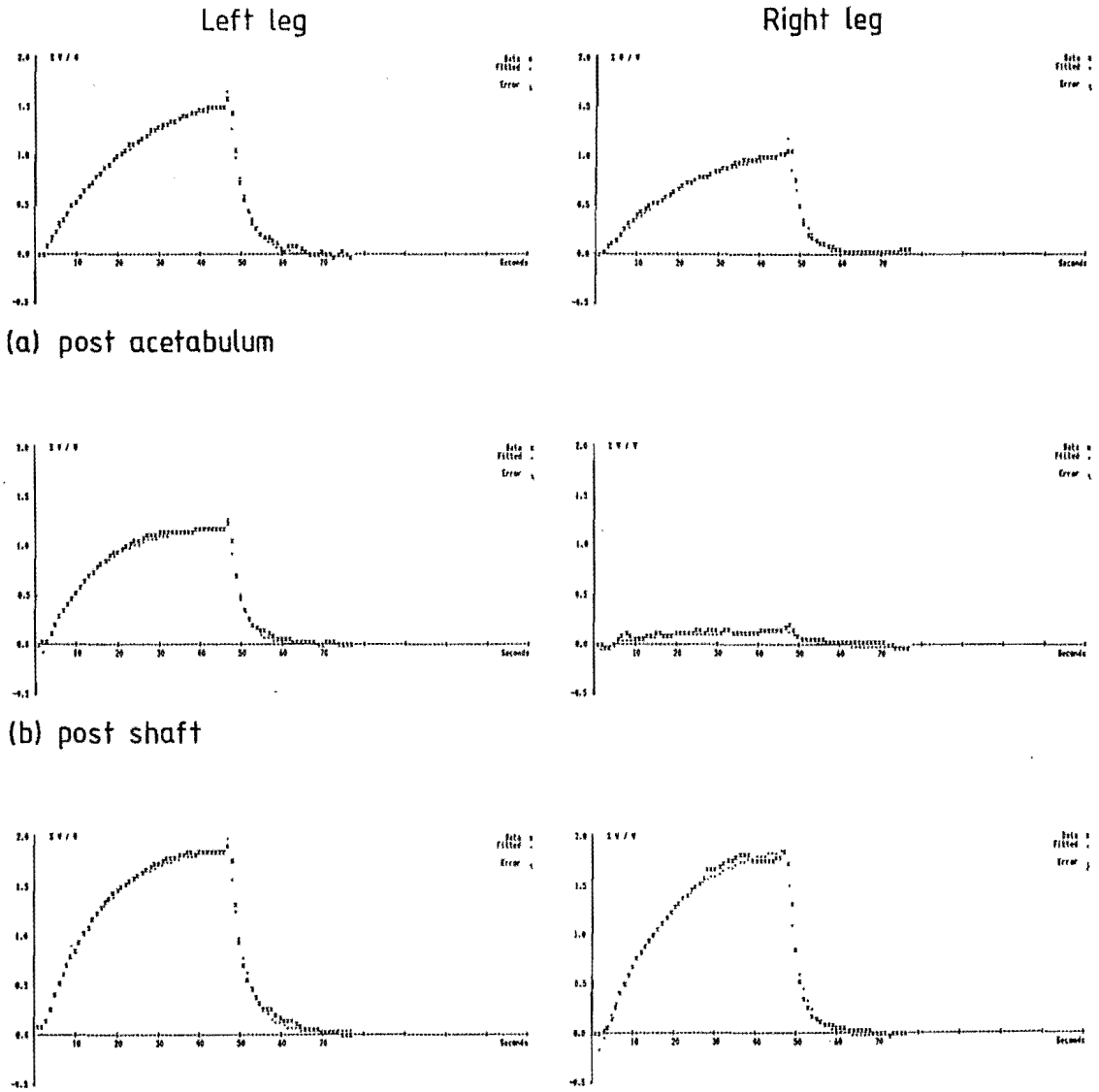


Figure 9.12: Volume change curves from venous occlusion plethysmography showing changes in leg blood flow during total hip replacement surgery. In each row the left hand plot is for the left leg and the right hand plot is for the right leg. The right leg was being operated on.
(a) Results from post acetabulum measurement.
(b) Results from post shaft measurement.
(c) Results from post relocation measurement.

leg was normal at the post acetabulum and post relocation measurements, but was almost completely absent at the post shaft measurement when the leg was twisted about. The measurements also show that the flow in the left (non-operative) leg remained normal throughout the operation and that when the right leg was relocated (Figure 9.12c) the blood flow in both legs increased above its values before the derangement of the right leg (Figure 9.12a). The increased blood flow after the right leg was relocated and the changes in left leg blood flow in response to changes in the right leg are indicative of the operation of body mechanisms controlling circulation.

9.6 DISCUSSION

Venous occlusion plethysmography provides a non-invasive means of measuring circulation in the limbs. The automated venous occlusion plethysmograph described in this chapter simplifies VOP by performing cuff inflation and deflation, measurement, and analysis automatically. The system produces printed output which provides a permanent record of the measurements for examination and archival storage.

The first clinical example presented in Section 9.5 (Figure 9.11) demonstrates the utility of the automated VOP in diagnosing deep vein thrombosis. From clinical use of the automated VOP and the earlier semi-manual VOP system, it has been found that venous occlusion plethysmography can reliably detect proximal DVT (i.e. those in the upper legs) but is less reliable for detecting DVT in the lower legs. This accords with the experience of others [Hull et al 1976, Hull et al 1978; see also Wheeler et al 1978]. The incidence of false positives is low, so that an abnormal curve of the type shown in Figure 9.11a almost always means that DVT is present.

The first example of Section 9.5 also demonstrates the utility of venous occlusion plethysmography as a non-invasive test of limb circulation in circumstances where the standard invasive tests (e.g. venography, in which a dye is injected into the foot and X-ray pictures are taken) are unsuitable. The patient on whom the measurements of Figure 9.11 were made was a young male adult admitted to the Intensive Care Unit (ICU) of Christchurch Hospital with spinal

injury from an accident resulting in partial paralysis. The prognosis with incomplete neurological injury (i.e. partial paralysis) is good provided the injury is not exacerbated by unnecessarily moving the patient. Therefore, instead of moving the patient from ICU to the Radiotherapy department for venography, the automated VOP was used in ICU as a non-invasive test. If the VOP output had been normal this would have been accepted as evidence that proximal DVT was not present and thus avoided the need for venography. However, the abnormal output shown in Figure 9.11 was obtained. As a result of this the patient underwent venography which confirmed the presence of proximal DVT. The patient was given anticoagulant drugs as treatment for the DVT.

Measurement of limb circulation is of physiological and clinical interest in a general sense, not merely for the diagnosis of deep vein thrombosis. For example, researchers at the Christchurch Clinical School of Medicine are interested in the changes of blood flow which occur as a result of surgery and anaesthesia. Venous occlusion plethysmography is a convenient technique which can be used intraoperatively. The second clinical example of Section 9.5 (Figure 9.12) illustrates the intraoperative information that the automated VOP can provide. Another study using the automated VOP which has recently begun at the Christchurch Clinical School is an investigation of how the blood flow in people's arms is affected by the posture of the arms. The automated venous occlusion plethysmograph is expected to have continuing application in Christchurch hospitals as a convenient non-invasive means of measuring limb circulation.

Part 3

IN CONCLUSION

CHAPTER 10

CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH

10.1 INTRODUCTION

Conclusions from the original work described in this thesis are presented in each of the relevant chapters, in Sections 5.3.6, 5.4.3, 6.6, 8.8, and 9.6. In this chapter the main points of the conclusions are reiterated, aspects of this work warranting further investigation are identified, and avenues for future research are proposed.

10.2 ASTRONOMICAL SPECKLE IMAGING

10.2.1 SHIFT-AND-ADD

Simple shift-and-add (ϕ3.6.2) and its extensions (ϕϕ 3.6.2, 3.6.3) are important members of the family of astronomical speckle processing techniques. The basic shift-and-add algorithms reviewed in Section 3.6.2 are computationally simpler to implement than any other speckle processing technique. Shift-and-add has produced useful astrophysical information [Bagnuolo and McAlister 1983, Hutter et al 1985, Baba et al 1985, Christou et al 1986]. The results presented in Chapter 5 of this thesis provide new understanding of the performance of simple shift-and-add when speckle images exhibit significant nonisoplanatism or significant photon-limiting. Both of these effects are likely to be often encountered in astronomical practice.

The results of Section 5.3 show that shift-and-add with partially isoplanatic speckle images experiences attenuation and loss of resolution of detail in the shift-and-add image, commeasurable with the similar effects which occur with speckle interferometry. No spurious fine detail is generated in the shift-and-add image as a result of partial isoplanatism. Shift-and-add is affected by

nonisoplanatism equivalently to speckle holography, and is affected more benignly than the Knox-Thompson method, in which the image is reportedly severely degraded by nonisoplanatism [Karo and Schneiderman 1975].

The empirical results of Section 5.3, which are drawn from computer simulations in which the atmosphere was modelled by a single phase screen, would be strengthened by modelling the atmosphere more realistically. The model of the atmosphere should be considerably improved by making the simulated atmospheric turbulence "thick" instead of "thin", i.e. by representing the atmosphere as two or possibly more phase screens (ϕ2.7) instead of just one. The phase screens could be simulated as described in Section 4.5 with different values of h. With a pair of phase screens, the two regions of large $C_N^2(h)$ at low and high altitudes (Figure 2.2) could be represented. Roddier et al's [1982] theoretical description of nonisoplanatic speckle images (ϕ2.6) has consequences for speckle interferometry that are different in some details for thick ($\Delta h \neq 0$ in (2.33)) and thin ($\Delta h = 0$) turbulence. The distinction between thick and thin turbulence is therefore important with speckle interferometry, and is probably also important with shift-and-add.

A theoretical description of the resolution loss suffered by shift-and-add with partial isoplanatism might be obtained by extending Hunt et al's [1983] analysis of shift-and-add to the nonisoplanatic case. Hunt et al derive expressions for the simple shift-and-add image. The expressions depend on the complex coherence factor [Goodman 1975 ϕ2.5] of the speckle images, which is in turn closely related to the ensemble average autocorrelation of the speckle images. The latter is the quantity evaluated in the image domain by speckle interferometry (3.14). The theoretical results of Roddier et al [1982] (5.25, 2.30) for speckle interferometry should therefore be applicable in generalising Hunt et al's theory for shift-and-add to embrace partial isoplanatism.

Shift-and-add processing of astronomical speckle images of the double star α^1, α^2 Scorpii (ϕ5.3.5) shows that the apparent angular diameter of α^1 Sco (the red supergiant Antares) is greater than that of α^2 Sco, despite the effects of nonisoplanatism and despite both

stars being below the Rayleigh limit of the 1.8m telescope used to form the speckle images. With a larger aperture, such as the 6.9m aperture of the whole Multiple Mirror Telescope [Hege et al 1985, cf. Christou et al 1986 Figure 7], the variant of shift-and-add discussed in Sections 5.3.4 and 5.3.5 should be able to form an image of α^1 Sco showing significant detail. This image, as well as being interesting in its own right, would serve as a useful comparison for images formed by other techniques such as weighted shift-and-add (ϕ3.6.3) and Fourier phase retrieval from speckle interferometry (ϕ3.4.2). All three types of image could be formed from a single data set, provided the data set is similar to that described in Section 5.3.5.1.

Section 5.4 presents results from a computational study of shift-and-add with photon-limited speckle images. The most important new result reported in this section is that the ghosting artefact of simple shift-and-add becomes more pronounced when the speckle images are severely photon-limited. The increase of ghosting seems to be associated with the ambiguity in determining the shift pixel for shift-and-add when speckle images have non-unique brightest pixels. The five shift strategies described in Section 5.4.2 represent preliminary steps in the search for shift strategies which are effective in photon-limited conditions. Shift strategies for non-unique brightest pixels should be comprehensively investigated. The correlation shift-and-add technique (ϕ3.6.3) may be applicable here.

Goodman and Belsher's [1976] analysis of speckle interferometry under photon-limited conditions should be applied to shift-and-add. This will not be a trivial exercise since doing so requires the probability density function of the position of the shift pixel to be expressed analytically. For practically useful shift strategies, i.e. those which choose pixels which are "brightest" or "nearly brightest" [cf. Bagnuolo 1982] in each speckle image, this pdf depends in a complicated manner upon the form of the object, the order statistics of the speckle image value, and the statistical properties of the seeing. Bagnuolo [1982, 1985a] eases this difficulty in his analyses of shift-and-add at high light levels by considering a shift strategy in which all pixels above a predefined fixed threshold of brightness, rather than a fixed number of brightest speckles in each image, are chosen as shift pixels. The same device might be helpful in adapting

Goodman and Belsher's analysis to shift-and-add. The results of such an analysis will be more strongly dependent on details of statistical models of atmospheric seeing [cf. Hunt et al 1983 ϕ6] than are the comparable results for speckle interferometry. They would, however, still represent an important and worthwhile extension to the understanding of shift-and-add.

10.2.2 ZERO-AND-ADD

Zero-and-add is a new speckle imaging technique that is fundamentally unrelated to almost all other speckle techniques so far proposed for optical astronomy (Chapter 3). Its principal advantage appears to be the possibility it offers of faithfully imaging arbitrary objects with no need to correct for the effects of atmospheric seeing, when applied to effectively isoplanatic speckle images. Its principal disadvantage is that the object must be of small extent, due to computational limits on the extents of the speckle images processed by zero-and-add. Several aspects of the computational studies of zero-and-add presented in Chapter 6 warrant further detailed investigation.

As yet there exists no comprehensive understanding of how various forms of contamination influence the quasi-steadfast zeros in the ZAA map, and particularly of how they cause migration of the zeros. The loci and distances of travel of the migration are likely to depend strongly on the detailed statistical characteristics of the seeing and the contamination. The fidelity of the zero-and-add image could be significantly improved if the migration could be even partly predicted. There is also a lack of understanding of how the form of the object and the statistics of the seeing influence the formation of obdurate zeros and the random background in the ZAA map. This understanding is required to optimally set the value of the threshold μ defined in Section 6.4.

In Chapter 6 the telescope point spread function is modelled as occupying only a single pixel of image space, which corresponds in practice to the pixel size of the detector being equal to the diffraction-limited resolution of the telescope. The effects on

zero-and-add of sampling image space more finely than the diffraction-limited resolution of the telescope should be studied. Following that, the effects of telescope aberrations on zero-and-add should be investigated. The former study could be readily begun by repeating the simulations reported in Chapter 6 with D set to (say) SIZE/2 or SIZE/4 (cf. §§ 4.3, 4.4). Note that the complex zeros of the telescope's transfer function $H_T(w)$ (cf. 2.8d) should not merely manifest themselves as steadfast zeros in the averaged zero map, even with ideal speckle images (6.20). This is because the speckle transfer function $H_m(w)$ is not separable into the product of $H_T(w)$ with a transfer function describing the atmosphere (§2.5). $H_T(w)$ might give rise to obdurate zeros in the averaged zero map. These ought to be suppressed in the same way as obdurate zeros due to the seeing are, by forming a ZAA map (§6.4).

It is important to devise means of ameliorating the contamination caused by truncating speckle images. Such truncation is often likely to be unavoidable in practice because of numerical limitations on the degree of polynomials which can be accurately factorised (cf. §6.2). Equally, improved means of factorising polynomials should be sought. The example presented in Section 6.5.3 suggests that rectangular windows cause quasi-steadfast zeros to migrate towards the real axis of the w-plane, Gaussian windows cause them to migrate away from the real axis, and combining rectangular and Gaussian windows cancels out the respective tendencies of the windows to cause migration. This result is, however, derived from one object viewed through one type of seeing. Further computational studies involving many different objects and seeing conditions are required to verify this result and develop optimum window functions for zero-and-add. The properties of windows at issue here are quite different from those of interest conventionally [Harris 1978].

Windowing is not the only means of ameliorating the effects of truncated images in image processing. Another approach, which has been used previously in the context of astronomical speckle imaging [Bates and Fright 1982 §§3.5,4] is edge extension [McDonnell and Bates 1975, Bates and McDonnell 198_ §15]. Edge extension removes the discontinuities at the edges of a truncated image by extrapolating the truncated image smoothly to zero outside its boundary. While edge

extension cannot be applied as straightforwardly to speckle images as can windowing, it might prove to be more effective than windowing for reducing the impact of truncation on zero-and-add. Algorithms for edge-extending speckle images, which are rapidly and randomly varying functions of position in image space, will have to be developed.

The zero-and-add algorithm defined in Section 6.4 does not recognise the possibility of two or more of an object's complex visibility zeros being effectively coincident. The algorithm could be extended to handle coincident zeros by examining the value of each local maximum chosen from the ZAA map in step (g), estimating thereby the multiplicity of the complex zero associated with the local maximum, and assigning a number of complex zeros to $\{\omega_p\}$ in step (h) equal to the estimated multiplicity. The probability of effectively coincident complex zeros of the object occurring should be investigated by comprehensive studies of different forms of object, with consideration given to the forms of object that are likely to be suitable for zero-and-add imaging in astronomical practice.

Chapter 6 has demonstrated the successful application of zero-and-add in digital computer simulations. An important next step is to test zero-and-add with speckle images formed in the optical laboratory. The experience gained from the optical laboratory will be valuable when applying zero-and-add to actual astronomical data.

10.3 PHYSIOLOGICAL MEASUREMENT IN MEDICINE

10.3.1 COMPUTERISED RESPIRATORY MONITOR

The respiratory monitoring system described in Chapter 8 performs continuous, breath-by-breath monitoring and analysis of breathing motions. Its non-invasive transducer causes minimal discomfort and inconvenience to the subject. This allows the system to be used for extended periods with hospital patients in clinical settings. The system incorporates a general-purpose microcomputer to perform on-line and off-line data processing. The system was funded

and built as a research tool for studying the respiratory effects of analgesic drugs administered to patients for pain relief during and after surgery. Its development in this role is not yet complete. However, it has already revealed significant differences in patients' breathing patterns before and after abdominal surgery when opioid narcotic drugs are used for pain relief. Preliminary results suggest that these changes do not occur when intercostal blocks (an alternative analgesic technique to opioid narcotics) are used.

Aspects of the respiratory monitoring system requiring further development are discussed in detail in Section 8.8. Here the two most important aspects, which affect the usefulness of the system in both the forthcoming clinical study of analgesic techniques and in the other possible applications mentioned in Section 8.8, are reiterated. The first aspect concerns the data analysis algorithms implemented on the system's microcomputer. At the moment, the system's data analysis software reveals long-term trends in breathing patterns but does not detect short-term irregularities of breathing. The latter type of breathing disturbance is of crucial importance in any type of apnoea study, for which the system is well suited. The second aspect requiring attention is the present lack of quantitative calibration of the system's non-invasive transducer for lung volume changes or respiratory gas flow. Such calibration would open up many applications in laboratory studies of pulmonary function, as the rapidly growing body of published results obtained with the commercial Respitrace system (cf. ϕ7.4.2) indicates. The issues involved in accurately calibrating the transducer in clinical settings are outlined in Sections 7.4.4, 7.5, and 8.8. The central issue, when calibrating the system's non-invasive transducer against a reference transducer of gas flow or volume, is to ensure that the breaths which are used in the calibration lead to as well-conditioned a solution as possible for the calibration coefficients (ϕ7.4.4). This is a challenging and important problem for future research. Another issue of significant but secondary importance is the development of optimum numerical techniques for effecting the solution. A systematic study of solution techniques has already begun [Suntheralingam 1985b] and this study should be continued.

10.3.2 AUTOMATED VENOUS OCCLUSION PLETHYSMOGRAPH

The automated venous occlusion plethysmograph described in Chapter 9 is now a complete, tested system. The system is mobile and easy to use. It makes venous occlusion plethysmography, a non-invasive test of limb blood circulation, convenient enough to be used in hospitals on a routine basis. It has already demonstrated its usefulness in Christchurch hospitals, both as a medical research tool and as an adjunct to established diagnostic tests in patient care. The system has been used to measure blood flow in the legs of patients undergoing total hip replacement surgery. It has also been used to non-invasively test for the presence of deep vein thrombosis, a relatively common complication of surgery, in patients admitted to hospital for surgery. These clinical applications, which are described in more detail in Section 9.5, are made possible by the non-invasive nature of venous occlusion plethysmography.

Possibilities for future research with the system exist in both medicine and medical engineering. Avenues of basic medical research in which venous occlusion plethysmography can play a major role include quantification of limb blood flow before, during, and after surgery, quantification of limb blood flow in patients with ischaemic vascular disease, and evaluation of the effects of vasoactive drugs on limb blood flow. Medical engineering research opportunities with the system centre on evaluating and improving the medical utility of the techniques embodied by the system: venous occlusion plethysmography, electrical impedance plethysmography, and the model of limb vasculature introduced by Seagar et al [1984]. An important area of research is to evaluate the potential of venous occlusion plethysmography for reliable, non-invasive screening for deep vein thrombosis of patients admitted to hospital for surgery.

REFERENCES

--- [1975] "Zooming in on Betelgeuse"
Scientific American, Vol. 232 No. 2, February. Pp. 42 - 43 (150)

--- [1978] "The prevention of postoperative venous thromboembolism" (editorial)
The New Zealand Medical Journal, Vol. 88 No. 622, 25 October.
Pp. 331 - 332 (349)

W.M. ABRAHAM, H. WATSON, A. SCHNEIDER, M. KING, L. YERGER, M.A. SACKNER [1981] "Noninvasive ventilatory monitoring by respiratory inductive plethysmography in conscious sheep"
Journal of Applied Physiology, Vol. 51 No. 6, December.
Pp. 1657 - 1661 (282, 289)

M. ABRAMOWITZ, I.A. STEGUN (ed.) [1973] "Handbook of Mathematical Functions"
Dover Publications, New York. (67, 179, 180)

N. AHMED, K.R. RAO [1975] "Orthogonal Transforms for Digital Signal Processing"
Springer-Verlag, Berlin. (171)

C. AIME [1974] "Measurement of averaged squared modulus of atmospheric-lens modulation transfer function"
Journal of the Optical Society of America, Vol. 64 No. 8, August.
Pp. 1129 - 1132 (66)

C. AIME, C. RODDIER, F. RODDIER [1975] "Holographic recording and reconstruction of astronomical images"
in "Imaging in Astronomy", Technical Digest from the Topical Meeting on Imaging in Astronomy, Cambridge, Massachusetts, 18 - 21 June.
Pp. ThC5-1 - ThC5-3 (124)

C. AIME, G. RICORT, J. HARVEY [1978a] "One-dimensional speckle interferometry of the solar granulation"
The Astrophysical Journal, Vol. 221 No. 1 Part 1, 1 April.
Pp. 362 - 367 (101)

C. AIME, G. RICORT, C. RODDIER, G. LAGO [1978b] "Changes in the atmospheric-lens modulation transfer function used for calibration in solar speckle interferometry"
Journal of the Optical Society of America, Vol. 68 No. 8, August.
Pp. 1063 - 1066 (101)

C. AIME, S. KADIRI, G. RICORT, C. RODDIER, J. VERNIN [1979] "Measurements of stellar speckle interferometry lens-atmosphere modulation transfer function"
Optica Acta, Vol. 26 No. 5, May. Pp. 575 - 581 (58, 71)

C. AIME, J. DEMARCQ, F. MARTIN, G. RICORT [1983] "One-dimensional telescope aperture for brightness and velocity speckle interferometry measurements"
Optical Engineering, Vol. 22 No. 2, March/April. Pp. 224 - 226 (265)

C. AIME, S. KADIRI, F. MARTIN, R. PETROV, G. RICORT [1984] "Measurement of submillisecond speckle displacements using a cross spectrum analysis technique. Test on atmospheric dispersion"
Astronomy and Astrophysics, Vol. 134 No. 2, May (II). Pp. 354 - 359 (90)

C. AIME, R.G. PETROV, F. MARTIN, G. RICORT, J. BORGNINO [1985] "Cross-spectrum techniques applied to astronomical speckle interferometry"
in "International Conference on Speckle", H.H. Arsenault (ed.), Proceedings of the SPIE, Vol. 556, 20 - 23 August. Pp. 297 - 310 (108)

G.J.M. AITKEN, D.L. DESAULNIERS [1979] "Restoration of atmospherically degraded images using complex spectral ratios"
Optics Communications, Vol. 28 No. 1, January. Pp. 26 - 29 (130)

G.J.M. AITKEN, R. HOUTMAN, R. JOHNSON, J.-M. POCHET [1985] "Direct phase gradient measurement for speckle image reconstruction"
Applied Optics, Vol. 24 No. 18, 15 September. Pp. 2926 - 2930 (124, 130, 131)

G.J.M. AITKEN, R. JOHNSON, R. HOUTMAN [1986] "Phase-gradient stellar image reconstruction"
Optics Communications, Vol. 56 No. 6, 15 January. Pp. 379 - 383 (131)

R.A. ALTES [1980] "Detection, estimation, and classification with spectrograms"
Journal of the Acoustical Society of America, Vol. 67 No. 4, April. Pp. 1232 - 1246 (336)

N. AMBROSINO, M. BRAMANTI [1983] "Non-invasive microwave monitoring of respiratory pattern"
Journal of Medical Engineering & Technology, Vol. 7 No. 5, September/October. Pp. 224 - 227 (279)

ANALOG DEVICES [1978] "RTI-1202 User's Reference Manual"
Analog Devices, Norwood, Massachusetts, 10 July. (356)

F.A. ANDERSON Jr, H.B. WHEELER [1979] "Venous occlusion plethysmography for the detection of venous thrombosis"
Medical Instrumentation, Vol. 13 No. 6, November-December. Pp. 350 - 354 (350)

F.A. ANDERSON Jr, B.C. PENNEY, N.A. PATWARDHAN, H.B. WHEELER [1980] "Impedance plethysmography: the origin of electrical impedance changes measured in the human calf"
Medical & Biological Engineering & Computing, Vol. 18 No. 2, March. Pp. 234 - 240 (350, 362)

J. ARSAC [1955] "Transmissions des fréquences spatiales dans les systèmes récepteurs d'ondes courtes"
Optica Acta, Vol. 2 No. 3, October. Pp. 112 - 118 (47)

N. BABA, F. KAWAGUCHI, T. OSE, S. ISOBE [1984] "Phase unwrapping method along radial coordinates for speckle image reconstruction"
Optics Communications, Vol. 49 No. 1, 1 February. Pp. 11 - 16 (99, 103, 122, 127, 128, 130)

N. BABA, S. ISOBE, Y. NORIMOTO, M. NOGUCHI [1985] "Stellar speckle image reconstruction by the shift-and-add method"
Applied Optics, Vol. 24 No. 10, 15 May. Pp. 1403 - 1405 (158, 377)

W.G. BAGNUOLO Jr [1982] "The application of Bates' algorithm to binary stars"
Monthly Notices of the Royal Astronomical Society, Vol. 200 No. 3, September. Pp. 1113 - 1122 (153, 157, 379)

W.G. BAGNUOLO Jr [1984] "Effect of nonisoplanaticity on the shift-and-add algorithm"
Optics Letters, Vol. 9 No. 3, March. Pp. 65 - 67 (78, 79, 81, 157, 189, 201, 216)

W.G. BAGNUOLO Jr [1985a] "Image restoration by the shift-and-add algorithm"
Optics Letters, Vol. 10 No. 5, May. Pp. 200 - 202 (156, 157, 158, 162, 163, 379)

W.G. BAGNUOLO Jr [1985b] "Image reconstruction via the shift-and-add algorithm"
in "Applications of Digital Image Processing VIII",
Proceedings of the SPIE, Vol. 575. Pp. 82 - 89 (157, 158, 162, 163)

W.G. BAGNUOLO Jr, H.A. McALISTER [1983] "The true nodal quadrant of Capella"
Publications of the Astronomical Society of the Pacific, Vol. 95 No. 574, December. Pp. 992 - 995 (158, 377)

G. BAIER, G. WEIGELT [1983] "Speckle interferometry observations of the asteroids Juno and Amphitrite"
Astronomy and Astrophysics, Vol. 121 No. 1, May (I). Pp. 137 - 141 (100)

G. BAIER, G. WEIGELT [1984] "Determination of the diameter of Pluto and its moon Charon"
Angewandte Optik, Annual Report, Physikalisches Institut der Universitat Erlangen-Nurnberg. P. 27 (100)

P.A. BAKUT, I.N. MATVEEV, K.N. SVIRIDOV, N.D. USTINOV, N.Yu. KHOMICH [1984] "Possibility of reconstructing the image of an object, undistorted by the atmosphere, from a single speckle interferogram"
Optics and Spectroscopy, Vol. 57 No. 1, July. Pp. 82 - 84 (151)

J.E. BALDWIN, P.J. WARNER [1978a] "Phaseless aperture synthesis"
Monthly Notices of the Royal Astronomical Society, Vol. 182 No. 2, February. Pp. 411 - 422 (115, 152)

J.E. BALDWIN, P.J. WARNER [1978b] "Fundamental aspects of aperture synthesis with limited or no phase information"
in "Image Formation from Coherence Functions in Astronomy", C. van Schooneveld (ed.), D. Reidel Publishing Company, Dordrecht, Holland, 1979. Proceedings of IAU Colloquium No. 49, Groningen, Netherlands, 10 - 12 August. Pp. 67 - 82 (115)

Yu.Yu. BALEGA, A. BLAZIT, D. BONNEAU, L. KOECHLIN, R. FOY, A. LABEYRIE [1982] "The angular diameter of Betelgeuse"
Astronomy and Astrophysics, Vol. 115 No. 2, November (II). Pp. 253 - 256 (100)

R. BARAKAT, P. NISENSON [1983] "The effectiveness of astronomical speckle transfer function reweighting algorithms"
Optics Communications, Vol. 45 No. 5, 1 May. Pp. 311 - 316 (102)

R. BARLETTI, G. CEPPATELLI, L. PATERNO, A. RIGHINI, N. SPERONI [1976] "Mean vertical profile of atmospheric turbulence relevant for astronomical seeing"
Journal of the Optical Society of America, Vol. 66 No. 12, December. Pp. 1380 - 1383 (54)

R. BARLETTI, G. CEPPATELLI, L. PATERNO, A. RIGHINI, N. SPERONI [1977] "Astronomical site testing with balloon borne radiosondes: results about atmospheric turbulence, solar seeing, and stellar scintillation"
Astronomy and Astrophysics, Vol. 54 No. 3, February (I). Pp. 649 - 659 (54, 61)

B.V. BARLOW [1975] "The Astronomical Telescope"
Wykeham Publications, London. (51, 53)

H. BARTELT, B. WIRNITZER [1985] "Shift-invariant imaging of photon-limited data using bispectral analysis"
Optics Communications, Vol. 53 No. 1, 1 February. Pp. 13 - 16 (133)

H. BARTELT, A.W. LOHMANN, B. WIRNITZER [1984] "Phase and amplitude recovery from bispectra"
Applied Optics, Vol. 23 No. 18, 15 September. Pp. 3121 - 3129 (131, 133, 135, 137, 138, 140, 141)

J. BARTHMAIER [1979] "Intel MULTIBUS Interfacing"
Intel application note AP-28A, Intel Corporation, Santa Clara, California. (356)

R.H.T. BATES [1969] "Contributions to the theory of intensity interferometry"
Monthly Notices of the Royal Astronomical Society, Vol. 142 No. 4. Pp. 413 - 428 (229, 233)

R.H.T. BATES [1971] "Holographic approach to radiation pattern measurement - I. General theory"
International Journal of Engineering Science, Vol. 9 No. 11, November. Pp. 1107 - 1121 (111)

R.H.T. BATES [1976] "A stochastic image restoration procedure"
Optics Communications, Vol. 19 No. 2, November. Pp. 240 - 244 (159)

R.H.T. BATES [1978] "On phase problems. II"
Optik, Vol. 51 No. 3, September. Pp. 223 - 234 (229)

R.H.T. BATES [1982a] "Fourier phase problems are uniquely solvable in more than one dimension. I: underlying theory"
Optik, Vol. 61 No. 3, June. Pp. 247 - 262 (118)

R.H.T. BATES [1982b] "Astronomical speckle imaging"
Physics Reports, Vol. 90 No. 4, October. Pp. 203 - 297 (54, 56, 57, 58, 64, 65, 66, 68, 79, 80, 82, 83, 84, 91, 92, 99, 102, 110, 117, 125, 128, 152, 153, 159, 176, 229)

R.H.T. BATES [1984] "Uniqueness of solutions to two-dimensional Fourier phase problems for localized and positive images"
Computer Vision, Graphics, and Image Processing, Vol. 25 No. 2, February. Pp. 205 - 217 (118)

R.H.T. BATES, F.M. CADY [1980] "Towards true imaging by wideband speckle interferometry"
Optics Communications, Vol. 32 No. 3, March. Pp. 365 - 369 (-14, 66, 145, 152, 153, 154, 170)

R.H.T. BATES, W.R. FRIGHT [1982] "Towards imaging with a speckle-interferometric optical synthesis telescope"
Monthly Notices of the Royal Astronomical Society, Vol. 198 No. 3, March. Pp. 1017 - 1031 (122, 130, 154, 159, 170, 229, 265, 381)

R.H.T. BATES, W.R. FRIGHT [1983] "Composite two-dimensional phase-restoration procedure"
Journal of the Optical Society of America, Vol. 73 No. 3, March. Pp. 358 - 365 (119, 121)

R.H.T. BATES, W.R. FRIGHT [1984] "Reconstructing images from their Fourier intensities"
Chapter 5 in "Image Reconstruction from Incomplete Observations", T.S. Huang (ed.), JAI Press, Greenwich, Connecticut. Advances in Computer Graphics and Image Processing, Vol. 1. Pp. 227 - 264 (117, 118, 119, 121, 229, 242)

R.H.T. BATES, P.T. GOUGH [1975] "New outlook on processing radiation received from objects viewed through randomly fluctuating media"
IEEE Transactions on Computers, Vol. C-24 No. 4, April. Pp. 449 - 456 (58, 129, 140)

R.H.T. BATES, M.J. McDONNELL [198_] "Image Restoration and Reconstruction"
Oxford University Press, (in press 1985). (83, 98, 156, 381)

R.H.T. BATES, M.O. MILNER [1978] "Towards imaging of star clusters by speckle interferometry"
in "Image Formation from Coherence Functions in Astronomy", C. van Schooneveld (ed.), D. Reidel Publishing Company, Dordrecht, Holland, 1979. Proceedings of IAU Colloquium No. 49, Groningen, Netherlands, 10 - 12 August. Pp. 187 - 193 (114, 151)

R.H.T. BATES, R.A. MINARD [1983] "Some new approaches to inverse scattering"
in "Inverse Optics", A.J. Devaney (ed.),
Proceedings of the SPIE, Vol. 413, 6 - 8 April. Pp. 56 - 60 (158)

R.H.T. BATES, R.A. MINARD [1984] "Compensation for multiple reflection"
IEEE Transactions on Sonics and Ultrasonics, Vol. SU-31 No. 4, July.
Pp. 330 - 336 (158)

R.H.T. BATES, P.J. NAPIER [1972] "Identification and removal of phase errors in interferometry"
Monthly Notices of the Royal Astronomical Society, Vol. 158 No. 4.
Pp. 405 - 424 (102, 229, 233)

R.H.T. BATES, F.L. NG [1972] "Polarisation-source formulation of electromagnetism and dielectric-loaded waveguides"
Proceedings of the IEE, Vol. 119 No. 11, November. Pp. 1568 - 1574
(29)

R.H.T. BATES, B.S. ROBINSON [1981] "Ultrasonic transmission speckle imaging"
Ultrasonic Imaging, Vol. 3 No. 4, October. Pp. 378 - 394 (68, 158)

R.H.T. BATES, B.S. ROBINSON [1982] "A stochastical imaging procedure"
Acoustical Imaging, Vol. 12, E.A. Ash, C.R. Hill (eds.), Plenum
Press, New York. Pp. 185 - 191 (158, 159)

R.H.T. BATES, D.G.H. TAN [1985] "Towards reconstructing phases of inverse-scattering signals"
Journal of the Optical Society of America A, Vol. 2 No. 11, November.
Pp. 2013 - 2018 (118)

R.H.T. BATES, P.T. GOUGH, P.J. NAPIER [1973] "Speckle interferometry gives holograms of multiple star systems"
Astronomy and Astrophysics, Vol. 22 no. 2, January (II).
Pp. 319 - 320 (112)

R.H.T. BATES, P.J. NAPIER, A.E. McKINNON, M.J. McDONNELL [1976a]
"Self-consistent deconvolution. I: theory"
Optik, Vol. 44 No. 2, January. Pp. 183 - 201 (229, 240, 253, 259)

R.H.T. BATES, W.M. BOERNER, G.R. DUNLOP [1976b] "An extended Rytov approximation and its significance for remote sensing and inverse scattering"
Optics Communications, Vol. 18 No. 4, September. Pp. 421 - 423 (34)

R.H.T. BATES, M.J. McDONNELL, P.T. GOUGH [1977] "Imaging through randomly fluctuating media"
Proceedings of the IEEE, Vol. 65 No. 1, January. Pp. 138 - 143 (150,
151)

R.H.T. BATES, R.M. LEWITT, M.J. McDONNELL, M.O. MILNER, T.M. PETERS
[1978a] "Practical Image Processing"
Physics in Technology, Vol. 9 No. 3, May. Pp. 101 - 107 (151)

R.H.T. BATES, M.O. MILNER, G.I. LUND, A.D. SEAGAR [1978b]
"Towards high resolution imaging by speckle interferometry"
Optics Communications, Vol. 26 No. 1, July. Pp. 22 - 26 (151)

R.H.T. BATES, W.R. FRIGHT, F.M. CADY, G.J. BERZINS [1981]
"Speckle processing, shift-and-add, and compensating for instrument aberrations"
in "Transformations in Optical Signal Processing", W.T. Rhodes, J.R. Fienup, B.E.A. Saleh (eds.), SPIE, Bellingham, Washington, 1984. Proceedings of the SPIE, Vol. 373, 23 - 25 February. Pp. 197 - 202 (154, 159)

R.H.T. BATES, B.R. HUNT, B.S. ROBINSON, W.R. FRIGHT, P.T. GOUGH [1982]
"Aspects of speckle interferometric imaging"
in "Electronic Image Processing",
Proceedings of International Conference, University of York,
26 - 28 July. IEE Conference Publication No. 214. Pp. 164 - 168 (158)

R.H.T. BATES, K.L. GARDEN, T.M. PETERS [1983] "Overview of computed tomography with emphasis on future developments"
Proceedings of the IEEE, Vol. 71 No. 3, March. Pp. 356 - 372 (29, 266)

J.H.T. BATES, W.R. FRIGHT, R.H.T. BATES [1984] "Wiener filtering and Cleaning in a general image processing context"
Monthly Notices of the Royal Astronomical Society", Vol. 211 No. 1, 1 November. Pp. 1 - 14 (156)

R.H.T. BATES, A.M. SINTON, R.A. MINARD [1985] "Generalization of shift-and-add imaging"
in "International Conference on Speckle", H.H. Arsenault (ed.), Proceedings of the SPIE, Vol. 556, 20 - 23 August. Pp. 263 - 269 (144, 160, 229)

R.H.T. BATES, J.C. CHRISTOU, W.J. COCKE, J.D. DRUMMOND, W.R. FRIGHT, E.K. HEGE, D. MNYAMA, F. RODDIER, C. RODDIER, A.M. SINTON [198_]
"Betelgeuse and Eros imaged by Fourier phase retrieval from speckle data"
In preparation, first draft May 1985. (121, 163)

J.M. BECKERS [1982] "Differential Speckle Interferometry"
Optica Acta, Vol. 29 No. 4, April. Pp. 361 - 362 (90)

J.M. BECKERS, E.K. HEGE [1981] "Experiments in differential speckle interferometry"
in "Instrumentation for Astronomy with Large Optical Telescopes", C.M. Humphries (ed.), D. Reidel Publishing Company, Dordrecht, Holland, 1982. Proceedings of IAU Colloquium No. 67, Zelenchukskaya, USSR, 8 - 10 September. Pp. 199 - 206 (90)

BECKMAN INSTRUMENTS [1965] "Biopotential Skin Electrode Instruction Manual"
Beckman Instruments Inc., Stanford Industrial Park, Palo Alto, California. (303)

W.R. BENNETT [1956] "Methods of solving noise problems"
Proceedings of the IRE, Vol. 44 No. 5, May. Pp. 609 - 638 (68)

M.J. BERAN, G.B. PARRENT Jr [1974] "Theory of Partial Coherence"
The Society of Photo-optical Instrumentation Engineers. (9, 12, 15, 16)

G.D. BERGLAND [1969] "A guided tour of the fast Fourier transform"
IEEE Spectrum, Vol. 6 No. 7, July. Pp. 41 - 52 (171, 336)

G.D. BERGLAND [1981] "A guided tour of program design methodologies"
Computer, Vol. 14 No. 10, October. Pp. 13 - 37 (319)

C.H. BEST, N.B. TAYLOR [1972] "The Living Body. A Text in Human Physiology", 4th edition
Chapman and Hall, London. (317)

G. BEYLKIN, M.L. ORISTAGLIO [1985] "Distorted-wave Born and distorted-wave Rytov approximations"
Optics Communications, Vol. 53 No. 4. Pp. 213 - 216 (30)

R.W. BOBERG [1980] "Proposed microcomputer system 796 bus standard"
Computer, Vol. 13 No. 10, October. Pp. 89 - 105 (356)

S. BOHME, W. FRICKE, H. HEFELE, I. HEINRICH, W. HOFMANN, D. KRAHN, V.R. MATAS, L.D. SCHMADEL, G. ZECH (eds.) [1984] "Astronomy and Astrophysics Abstracts",
Vol. 35/36, Author and Subject indexes to Vols. 25-34, Literature 1979 - 1983. Springer-Verlag, Berlin. (99)

G. BONNEAU, M. BERTRAND, R. GUARDO, M. LAFORTUNE, A.R. LeBLANC [1983] "A preprocessor for on-line data reduction in accelerated playback of ECG's"
IEEE Transactions on Biomedical Engineering, Vol. BME-30 No. 3, March. Pp. 197 - 200 (346)

M. BORN, E. WOLF [1975] "Principles of Optics", 5th edition
Pergamon Press, Oxford. (9, 11, 13, 16, 17, 19, 20, 21, 41, 43, 46, 48, 84)

R.N. BRACEWELL [1978] "The Fourier Transform and its Applications", 2nd edition
McGraw-Hill, New York. (-9, -8, -7, 4, 6, 8, 9, 20, 21, 49, 94, 128, 130, 171, 174, 196, 322, 326)

R.N. BRACEWELL [1979] "Computer image processing"
in "Annual Review of Astronomy and Astrophysics", Vol. 17, Annual Reviews Inc., Palo Alto, California. Pp. 113 - 134 (91)

B.J. BRAMES, J.C. DAINTY [1981] "Method for determining object intensity distributions in stellar speckle interferometry"
Journal of the Optical Society of America, Vol. 71 No. 12, December. Pp. 1542 - 1545 (130, 145)

J.B. BRECKINRIDGE [1976] "Measurement of the amplitude of phase excursions in the earth's atmosphere"
Journal of the Optical Society of America, Vol. 66 No. 2, February. Pp. 143 - 144 (56)

J.B. BRECKINRIDGE [1978] "A white-light amplitude interferometer with 180-degree rotational shear"
Optical Engineering, Vol. 17 no. 2, March-April. Pp. 156 - 159 (85)

E.D. BRIGHAM [1974] "The Fast Fourier Transform"
Prentice-Hall, Englewood Cliffs, New Jersey. (7, 171)

D.S. BROWN, R.J. SCADDAN [1979] "An interferometer for efficient measurement of atmosphere MTF"
The Observatory, Vol. 99 No. 1031, August. Pp. 125 - 128 (61, 85)

Yu.M. BRUCK, L.G. SODIN [1979] "On the ambiguity of the image reconstruction problem"
Optics Communications, Vol. 30 No. 3, September. Pp. 304 - 308 (118, 233)

Yu.M. BRUCK, L.G. SODIN [1980] "A method for processing speckle images requiring no reference point source"
Astronomy and Astrophysics, Vol. 87 No. 1/2, July (I). Pp. 188 - 191 (102, 103)

Yu.M. BRUCK, L.G. SODIN [1984] "Speckle interferometry image reconstruction from the Fourier transform phase"
Journal of the Optical Society of America A, Vol. 1 No. 1, January. Pp. 73 - 80 (130)

J. BRYDON [1976] "Automatic monitoring of cardiac arrhythmias"
Chapter 2 in "IEE Medical Electronics Monographs 18 - 22", D.W. Hill, B.W. Watson (eds.), Peter Peregrinus. Pp. 27 - 41 (346)

M.J. BUERGER [1959] "Vector Space and its Applications in Crystal-Structure Investigation"
John Wiley & Sons, New York. (114, 115)

F.M. CADY [1980] "Applications of Microcomputers in Interactive Image Processing"
Ph.D. thesis, University of Canterbury, Christchurch. (153, 154)

F.M. CADY, R.H.T. BATES [1980] "Speckle processing gives diffraction-limited true images from severely aberrated instruments"
Optics Letters, Vol. 5 No. 10, October. Pp. 438 - 440 (153, 154)

J.S. CALNAN, F. ALLENBY [1975] "The prevention of deep vein thrombosis after surgery"
British Journal of Anaesthesia, Vol. 47 No. 2, February. Pp. 151 - 160 (349, 358)

K.R. CASTLEMAN [1979] "Digital Image Processing"
Prentice-Hall, Englewood Cliffs, New Jersey. (98, 210)

D.M. CATLEY, C. THORNTON, C. JORDAN, J.R. LEHANE, D. ROYSTON, J.G. JONES [1982a] "Postoperative hypoxaemia and its association with respiratory pattern and analgesic regime" (conference abstract)
Federation Proceedings, Vol. 41 No. 4, 5 March. P. 1378 (282, 295, 296)

D.M. CATLEY, C. THORNTON, C. JORDAN, D. ROYSTON, J.R. LEHANE, J.G. JONES [1982b] "Continuous postoperative monitoring reveals oxygen desaturation associated with paradoxical respiration" (conference abstract)
American Review of Respiratory Disease (annual meeting supplement), Vol. 125 No. 4 Part 2, April. P. 105 (282, 290, 295, 296)

J.A. CATLING, D.M. PINTO, C. JORDAN, J.G. JONES [1980] "Respiratory effects of analgesia after cholecystectomy: comparison of continuous and intermittent papaveretum"
British Medical Journal, Vol. 281 No. 6238, 16 August. Pp. 478 - 480 (282, 296)

T.S. CHADHA, A.W. SCHNEIDER, M.A. SACKNER [1981] "Effect of minimal bronchoconstriction on ventilatory pattern in normal and asthmatic subjects" (conference abstract)
American Review of Respiratory Disease (annual meeting supplement), Vol. 123 No. 4 Part 2, April. P. 205 (282)

T.S. CHADHA, H. WATSON, S. BIRCH, G.A. JENOURI, A.W. SCHNEIDER, M.A. COHN, M.A. SACKNER [1982] "Validation of respiratory inductive plethysmography using different calibration procedures"
American Review of Respiratory Disease, Vol. 125 No. 6, June. Pp. 644 - 649 (282, 285, 289)

A. CHELLI, P. LENA, C. RODDIER, F. RODDIER, F. SIBILLE [1979] "Modulation transfer function for infra-red stellar speckle interferometry: evidence for a log-normal statistic"
Optica Acta, Vol. 26 No. 5, May. Pp. 583 - 595 (58, 71)

A. CHELLI, C. PERRIER, Y.G. BIRAUD [1983] "One-dimensional high resolution image reconstruction on Eta Carinae at 4.6μm with speckle data"
Astronomy and Astrophysics, Vol. 117 No. 2, January (II). Pp. 199 - 204 (129)

L.A. CHERNOV [1967] "Wave Propagation in a Random Medium" (translated from the Russian)
Dover Publications, New York. (32, 33)

S.D. CHOI, B.C. SPAULDING, J.B. GROSS, J.L. APFELBAUM [1985] "Comparison of the ventilatory effects of etomidate and methohexital"
Anesthesiology, Vol. 62 No. 4, April. Pp. 442 - 447 (296)

J.C. CHRISTOU [1981] "Imaging of star clusters from speckle interferometry"
Optics Communications, Vol. 37 No. 5, 1 June. Pp. 331 - 334 (115, 117)

J.C. CHRISTOU [1985a] "Seeing Effects and their Calibration for Astronomical Speckle Interferometry Observations"
Ph.D. thesis draft, New Mexico State University, Las Cruces, New Mexico. (100, 163, 164, 209)

J.C. CHRISTOU [1985b] Private communication to R.H.T. Bates
Steward Observatory, University of Arizona, Tucson, 15 February. (216)

J.C. CHRISTOU, E.K. HEGE, J. FREEMAN, P.A. STRITTMATTER [1984] "Speckle image reconstruction: weighted shift-and-add analysis" (conference abstract)
Bulletin of the American Astronomical Society, Vol. 16 No. 4. P. 885 (163)

J.C. CHRISTOU, E.K. HEGE, J.D. FREEMAN, E. RIBAK [1985] "Images from astronomical speckle data: weighted shift-and-add analysis"
in "International Conference on Speckle", H.H. Arsenault (ed.), Proceedings of the SPIE, Vol. 556, 20 - 23 August. Pp. 255 - 262 (163, 164, 165, 218, 222, 227)

J.C. CHRISTOU, E.K. HEGE, J.D. FREEMAN, E. RIBAK [1986] "A self calibrating shift-and-add technique for speckle imaging"
Journal of the Optical Society of America A, Vol. 3 No. 2, February. Pp. 204 - 209 (163, 164, 216, 222, 227, 377, 379)

J.C. CHRISTOU, E. RIBAK, E.K. HEGE, J.D. FREEMAN [198_] "Images from astronomical speckle data: weighted shift-and-add analysis"
Optical Engineering, submitted for publication 1985. (163, 164)

J.J. CHURCH [1979] "Continuous narcotic infusions for relief of postoperative pain"
British Medical Journal, Vol. 1 No. 6169, 14 April. Pp. 977 - 979 (295)

F. CLERGUE, C. MONTEMBAULT, O. DESPIERRES, F. GHESQUIERE, A. HARARI, P. VIARS [1984] "Respiratory effects of intrathecal morphine after upper abdominal surgery"
Anesthesiology, Vol. 61 No. 6, December. Pp. 677 - 685 (295)

W.J. COCKE [1980] "Computer simulation comparisons of speckle image reconstruction techniques"
Proceedings of the SPIE, Vol. 231. Pp. 99 - 105 (122, 123, 124, 130)

J.D. COHEN [1985] "High bandwidth triple product processor using a shearing interferometer"
Applied Optics, Vol. 24 No. 19, 1 October. Pp. 3173 - 3178 (140)

M.A. COHN, H. WATSON, R. WEISSHAUT, F. STOTT, M.A. SACKNER [1978] "A transducer for non-invasive monitoring of respiration"
in "Proceedings of the Second International Symposium on Ambulatory Monitoring (ISAM 1977)", F.D. Stott, E.B. Raftery, P. Sleight, L. Goulding (eds), Academic Press, London. Pp. 119 - 128 (not sighted) (280, 281)

M.A. COHN, A.S.V. RAO, M. BROUDY, S. BIRCH, H. WATSON, N. ATKINS, B. DAVIS, F.D. STOTT, M.A. SACKNER [1982] "The respiratory inductive plethysmograph: a new non-invasive monitor of respiration"
Bulletin Europeen de Physiopathologie Respiratoire, Vol. 18 No. 4, July-August. Pp. 643 - 658 (not sighted) (281)

J.W. COOLEY, J.W. TUKEY [1965] "An algorithm for the machine calculation of complex Fourier series"
Mathematics of Computation, Vol. 19 No. 90, April. Pp. 297 - 301 (171)

L.A. COX, A.P. ALMEIDA, J.S. ROBINSON, J.K. HORSLEY [1974] "An electronic respirometer"
British Journal of Anaesthesia, Vol. 46 No. 4, April. Pp. 302 - 310 (277)

R.A. CRANE, B. STUTTARD [1976] "A digital technique for linearising the output of a turbine anenometer"
Biomedical Engineering, Vol 11 No. 1, January. Pp. 19 - 21 (277)

D.W. CUTTER, A.W. LOHMANN [1974] "Achromatic stellar interferometry"
Optics Communications, Vol. 12 No. 2, October. Pp. 220 - 222 (83)

J.C. DAINTY [1973] "Diffraction-limited imaging of stellar objects using telescopes of low optical quality"
Optics Communications, Vol. 7 No. 2, February. Pp. 129 - 134 (58, 66, 70, 71)

J.C. DAINTY [1974] "The transfer function, signal-to-noise ratio, and limiting magnitude in stellar speckle interferometry"
Monthly Notices of the Royal Astronomical Society, Vol. 169 No. 3, December. Pp. 631 - 641 (58, 71)

J.C. DAINTY (ed.) [1975a] "Laser Speckle and Related Phenomena"
Springer-Verlag, Berlin. (64, 67)

J.C. DAINTY [1975b] "Stellar speckle interferometry"
in "Laser Speckle and Related Phenomena", J.C. Dainty (ed.), Springer-Verlag, Berlin. Pp. 255 - 280 (64, 71, 91, 150)

J.C. DAINTY [1977] "Telescope requirements for speckle interferometry"
in "Optical Telescopes of the Future", F. Pacini, W. Richter, R.N. Wilson (eds.),
Proceedings of ESO Conference, Geneva, 12 - 15 December.
Pp. 43 - 46 (71)

J.C. DAINTY [1984] "Stellar speckle interferometry"
in "Laser Speckle and Related Phenomena", 2nd edition, J.C. Dainty (ed.), Springer-Verlag, Berlin. Pp. 255 - 320 (64, 66, 91, 99, 100, 106, 108, 109, 117, 123, 124, 126, 129, 141, 145, 150, 169, 229)

J.C. DAINTY, M.A. FIDDY [1984] "The essential role of prior knowledge in phase retrieval"
Optica Acta, Vol. 31 No. 3, March. Pp. 325 - 330 (118)

J.C. DAINTY, R.J. SCADDAN [1974] "A coherence interferometer for direct measurement of the atmospheric transfer function"
Monthly Notices of the Royal Astronomical Society, Vol. 167 No. 3, June. Pp. 69P - 73P (62, 85)

J.C. DAINTY, R. SHAW [1974] "Image Science (Principles, Analysis, and Evaluation of Photographic-type Imaging Processes)"
Academic Press, London. (105, 179)

B.L.K. DAVEY, A.M. SINTON, R.H.T. BATES [198_] "Zero-and-add"
Optical Engineering, accepted for publication 16 January 1986. (144, 229)

D.H. DE VORKIN [1975] "Michelson and the problem of stellar diameters"
Journal of the History of Astronomy, Vol. 6 No. ?, ?. Pp. 1 - 18 (83, 84)

DEC [1984] "Programming in VAX FORTRAN"
Manual No. AA-D034D-TE, Digital Equipment Corporation, Maynard, Massachusetts, September. (175)

H.R. DELGADO, S.R. BRAUN, J.B. SKATRUD, W.G. REDDAN, D.F. PEGELOW [1982] "Chest wall and abdominal motion during exercise in patients with chronic obstructive pulmonary disease"
American Review of Respiratory Disease, Vol. 126 No. 2, August. Pp. 200 - 205 (282)

D.R. DERBYSHIRE, A. BELL, P.A. PARRY, G. SMITH [1985] "Morphine sulphate slow release. Comparison with i.m. morphine for postoperative analgesia"
British Journal of Anaesthesia, Vol. 57 No. 9, September. Pp. 858 - 865 (295)

R. DERON, J.C. FONTANELLA [1984] "Reconstruction of atmospheric turbulence degraded images using the Knox and Thompson algorithm. Theoretical and experimental study" (in French)
Journal of Optics (Paris), Vol. 15 No. 1, January-February. Pp. 15 - 23 (not sighted) (129)

DIGITAL RESEARCH [1981] "Pascal/MT+ Language Reference Manual Release 5"
Digital Research, Pacific Grove, California. (318)

DIGITAL RESEARCH [1982] "SPP SpeedProgramming Package User's Guide Release 5.2"
Digital Research, Pacific Grove, California. (319)

T. DOLFIN, P. DUFFTY, D.L. WILKES, M.H. BRYAN [1982] "Calibration of respiratory induction plethysmography (Respitrace) in infants"
American Review of Respiratory Disease, Vol. 126 No. 3, September. Pp. 577 - 579 (282, 289)

T. DOLFIN, P. DUFFTY, D. WILKES, S. ENGLAND, H. BRYAN [1983] "Effects of a face mask and pneumotachograph on breathing in sleeping infants"
American Review of Respiratory Disease, Vol. 128 No. 6, December. Pp. 977 - 979 (278, 282)

R.D. DRIPPS, J.E. ECKENHOFF, L.D. VANDAM (eds.) [1982] "Introduction to Anesthesia. The Principles of Safe Practice", 6th edition
W.B. Saunders, Philadelphia. (294, 295, 339)

J.D. DRUMMOND, W.J. COCKE, E.K. HEGE, P.A. STRITTMATTER, J.V. LAMBERT [1985a] "Speckle interferometry of asteroids. I. 433 Eros"
Icarus, Vol. 61 No. 1, January. Pp. 132 - 151 (100)

J.D. DRUMMOND, E.K. HEGE, W.J. COCKE, J.D. FREEMAN, J.C. CHRISTOU, R.P. BINZEL [1985b]
"Speckle interferometry of asteroids. II. 532 Herculina"
Icarus, Vol. 61 No. 2, February. Pp. 232 - 240 (100)

P. DUFFTY, L. SPRIET, M.H. BRYAN, A.C. BRYAN [1981] "Respiratory induction plethysmography (Respitrace): an evaluation of its use in the infant"
American Review of Respiratory Disease, Vol. 123 No. 5, May.
Pp. 542 - 546 (282, 289)

G.R. DUNLOP, W.M. BOERNER, R.H.T. BATES [1976] "On an extended Rytov approximation and its comparison with the Born approximation"
Digest of papers, IEEE AP-S International Symposium 1976, University of Massachusetts, Amherst, Massachusetts, 11 - 15 November. IEEE catalogue number 76 CH1121-3AP. Pp. 587 - 591 (35)

H.M. DYCK, R.R. HOWELL [1983] "Seeing measurements at Mauna Kea from infrared speckle interferometry"
Publications of the Astronomical Society of the Pacific, Vol. 95 No. 572, October. Pp. 786 - 791 (62, 63)

H.M. DYCK, R.R. HOWELL [1985] "Astronomical speckle interferometry in the infra-red"
in "International Conference on Speckle", H.H. Arsenault (ed.), Proceedings of the SPIE, Vol. 556, 20 - 23 August. Pp. 274 - 278 (100, 169)

J. EBERSBERGER, G.P. WEIGELT [1979] "Speckle interferometry and speckle holography with the 1.5m and 3.6m ESO telescopes"
The Messenger, No. 18, September. Pp. 24 - 27 (112)

J. EBERSBERGER, G. WEIGELT [1985] "Isoplanicity and lifetime of stellar speckle interferograms"
Optica Acta, Vol. 32 No. 7, July. Pp. 793 - 801 (66, 79, 217)

M.J. ECCLES, M.E. SIM, K.P. TRITTON [1983] "Low Light Level Detectors in Astronomy"
Cambridge University Press, Cambridge. (51)

D.C. EHN, P. NISENSON [1975] "Astronomical speckle imaging" (conference abstract)
Journal of the Optical Society of America, Vol. 65 No. 10, October. P. 1196 (130)

S.E. ELLIOTT, J.H. SHORE, C.W. BARNES, J.LINDAUER, J.J. OSBORN [1977] "Turbulent airflow meter for long-term monitoring in patient-ventilator circuits"
Journal of Applied Physiology, Vol. 42 No. 3, March. Pp. 456 - 460 (277)

K.A. ELMQUIST, H. FULLMER, D.B. GUSTAVSON, G. MORROW [1979] "Standard specification for S-100 bus interface devices"
Computer, Vol. 12 No. 7, July. Pp. 28 - 52 (318)

A. ENGLANDER, M. SLATKINE, R. KAROUBI, D. BENSIMON [1983] "Probabilistic diffraction limited imaging through turbulence"
Optical Engineering, Vol. 22 No. 1, January/February. Pp. 145 - 148 (82)

A.E. ENNOS [1984] "Speckle interferometry"
Chapter 6 in "Laser Speckle and Related Phenomena", 2nd edition, J.C. Dainty (ed.), Springer-Verlag, Berlin. Pp. ? (not sighted) (90)

R.K. ERF (ed.) [1978] "Speckle Metrology"
Academic Press, New York. (90)

D. FAITHFULL, J.G. JONES, C. JORDAN [1979] "Measurement of the relative contributions of rib cage and abdomen/diaphragm to tidal breathing in man"
British Journal of Anaesthesia, Vol. 51 No. 5, May. Pp. 391 - 398 (280)

R.L. FANTE [1979] "Comments on a method for processing stellar speckle data"
Journal of the Optical Society of America, Vol. 69 No. 10, October. Pp. 1394 - 1396 (102)

L.B. FELSEN, N. MARCUVITZ [1973] "Radiation and Scattering of Waves"
Prentice-Hall, Englewood Cliffs, New Jersey. (34)

J.R. FIENUP [1978] "Reconstruction of an object from the modulus of its Fourier transform"
Optics Letters, Vol. 3 No. 1, July. Pp. 27 - 29 (119)

J.R. FIENUP [1979] "Space object imaging through the turbulent atmosphere"
Optical Engineering, Vol. 18 No. 5, September/October. Pp. 529 - 534 (119)

J.R. FIENUP [1981a] "Reconstruction and synthesis applications of an iterative approach"
in "Transformations in Optical Signal Processing", W.T. Rhodes, J.R. Fienup, B.E.A. Saleh (eds.), SPIE, Bellingham, Washington, 1984. Proceedings of the SPIE, Vol. 373, 23 - 25 February. Pp. 147 - 160 (119)

J.R. FIENUP [1981b] "Autocorrelation unfolding"
in "Transformations in Optical Signal Processing", W.T. Rhodes, J.R. Fienup, B.E.A. Saleh (eds.), SPIE, Bellingham, Washington, 1984. Proceedings of the SPIE, Vol. 373, 23 - 25 February. Pp. 203 - 209 (115, 117)

J.R. FIENUP [1982] "Phase retrieval algorithms: a comparison"
Applied Optics, Vol. 21 No. 15, 1 August. Pp. 2758 - 2769 (119)

J.R. FIENUP [1983] "Reconstruction of objects having latent reference points"
Journal of the Optical Society of America, Vol. 73 No. 11, November. Pp. 1421 - 1426 (117, 118)

J.R. FIENUP [1984] "Comparison of phase retrieval algorithms"
Chapter 4 in "Image Reconstruction from Incomplete Observations", T.S. Huang (ed.), JAI Press, Greenwich, Connecticut. Advances in Computer Graphics and Image Processing, Vol. 1. Pp. 191 - 225 (119, 120, 122)

J.R. FIENUP, G.B. FELDKAMP [1980] "Astronomical imaging by processing stellar speckle interferometry data"
in "Applications of Speckle Phenomena", W.H. Carter (ed.),
Proceedings of the SPIE, Vol. 243, 29 - 30 July. Pp. 95 - 102 (119)

J.R. FIENUP, T.R. CRIMMINS, W. HOLSZTYNSKI [1982] "Reconstruction of the support of an object from the support of its autocorrelation"
Journal of the Optical Society of America, Vol. 72 No. 5, May.
Pp. 610 - 624 (115, 117)

A. FLEISCH [1925] "Der Pneumotachograph - ein Apparat zur Geschwindigkeitsregistrierung der Atemluft"
Pflügers Archiv fur die Gesamte Physiologie des Menschen und der Tiere, Vol. 209 No. ?, ?. Pp. 713 - 722 (not sighted) (276)

C.I. FRANKS, B.H. BROWN, D.M. JOHNSTON [1976] "Contactless respiration monitoring of infants"
Medical and Biological Engineering, Vol. 14 No. 3, May.
Pp. 306 - 312 (278, 279)

J.D. FREEMAN, E. RIBAK, J.C. CHRISTOU, E.K. HEGE [1985] "Statistical analysis of the weighted shift-and-add image reconstruction technique"
in "International Conference on Speckle", H.H. Arsenault (ed.),
Proceedings of the SPIE, Vol. 556, 20 - 23 August. Pp. 279 - 283
(157, 165)

D.L. FRIED [1966] "Optical resolution through a randomly inhomogeneous medium for very long and very short exposures"
Journal of the Optical Society of America
Vol. 56 No. 10, October. Pp. 1372 - 1379 (48, 56, 60, 61, 82, 101)

D. L. FRIED [1976] "Varieties of isoplanatism"
in "Imaging Through the Atmosphere",
Proceedings of the SPIE, Vol. 75, 22 - 23 March. Pp. 20 - 29 (73, 74, 75)

D.L. FRIED [1978] "Probability of getting a lucky short-exposure image through turbulence"
Journal of the Optical Society of America, Vol. 68 No. 12, December.
Pp. 1651 - 1658 (82)

D.L. FRIED [1979] "Angular dependence of the atmospheric turbulence effect in speckle interferometry"
Optica Acta, Vol. 26 No. 5, May. Pp. 597 - 613 (71, 74, 75, 128, 129, 217)

D.L. FRIED [1982] "Anisoplanatism in adaptive optics"
Journal of the Optical Society of America, Vol. 72 No. 1, January.
Pp. 52 - 61 (75, 83)

J.P. FRIEL (ed.) [1974] "Dorland's Illustrated Medical Dictionary", 25th edition
W.B. Saunders, Philadelphia. (-5)

W.R. FRIGHT [1984] "The Fourier Phase Problem"
Ph.D. thesis, University of Canterbury, Christchurch. (117, 118, 121, 229, 233)

R.L. FROST, C.K. RUSHFORTH, B.S. BAXTER [1979] "Fast FFT-based algorithm for phase estimation in speckle imaging"
Applied Optics, Vol. 18 No. 12, 15 June. Pp. 2056 - 2061 (129)

L.A. GEDDES, L.E. BAKER [1975] "Principles of Applied Biomedical Instrumentation", 2nd edition
John Wiley & Sons, New York. (280)

R.W. GERCHBERG, W.O. SAXTON [1972] "A practical algorithm for the determination of phase from image and diffraction plane pictures"
Optik, Vol. 35 No. 2, April. Pp. 237 - 246 (119)

D.Y. GEZARI, A. LABEYRIE, R.V. STACHNIK [1972] "Speckle interferometry: diffraction-limited measurements of nine stars with the 200-inch telescope"
The Astrophysical Journal, Vol. 173 No. 1 Part 2, 1 April.
Pp. L1 - L5, Plates L1, L2 (99)

R. GILBERT, J.H. AUCHINCLOSS Jr, J. BRODSKY, W. BODEN [1972] "Changes in tidal volume, frequency, and ventilation induced by their measurement"
Journal of Applied Physiology, Vol. 33 No. 2, August. Pp. 252 - 254 (278)

J. GOLDING, S. LIMERICK, A. MACFARLANE [1985] "Sudden Infant Death. Patterns, Puzzles and Problems"
Open Books, Near Shepton Mallet, Somerset. (346)

H. GONZALEZ, B. HALLER, H.L. WATSON, M.A. SACKNER [1984] "Accuracy of respiratory inductive plethysmograph over wide range of rib cage and abdominal compartmental contributions to tidal volume in normal subjects and in patients with chronic obstructive pulmonary disease"
American Review of Respiratory Disease, Vol. 130 No. 2, August.
Pp. 171 - 174 (282, 289, 290)

R.C. GOODE, E.B. BROWN Jr, M.G. HOWSON, D.J.C. CUNNINGHAM [1969] "Respiratory effects of breathing down a tube"
Respiration Physiology, Vol. 6 No. 3, April. Pp. 343 - 359 (273)

R. GOODISON [1983] "Epidural Narcotics"
Anaesthesia and Intensive Care, Vol. 11 No. 4, November. P. 389 (295)

J.W. GOODMAN [1968] "Introduction to Fourier Optics"
McGraw-Hill, San Francisco. (-9, -8, -7, 24, 35, 38, 39, 41, 43, 46, 47, 49)

J.W. GOODMAN [1970] "Analogy between holography and interferometric image formation"
Journal of the Optical Society of America, Vol. 60 No. 4, April.
Pp. 506 - 509 (112)

J.W. GOODMAN [1975] "Statistical properties of laser speckle patterns"
in "Laser Speckle and Related Phenomena", J.C. Dainty (ed.),
Springer-Verlag, Berlin. Pp. 9 - 75 (67, 68, 69, 72, 378)

J.W. GOODMAN, J.F. BELSHER [1976] "Fundamental limitations in linear invariant restoration of atmospherically degraded images" in "Imaging Through the Atmosphere", Proceedings of the SPIE, Vol. 75, 22 - 23 March. Pp. 141 - 154 (107, 129, 141, 179, 379)

A. GOODMAN GILMAN, L.S. GOODMAN, A. GILMAN (eds.) [1980] "The Pharmacological Basis of Therapeutics", 6th edition Macmillan, New York. (294)

P.T. GOUGH, R.H.T. BATES [1974] "Speckle holography" Optica Acta, Vol. 21 No. 3, March. Pp. 243 - 254 (112)

D.J. GRANRATH [1984] "Maximum-magnitude estimation of the object's power spectrum in stellar speckle interferometry" Optics Letters, Vol. 9 No. 11, November. Pp. 478 - 480 (103, 105)

A.H. GREENAWAY [1982] "Diffraction-limited pictures from single turbulence-degraded images in astronomy" Optics Communications, Vol. 42 No. 3, 1 July. Pp. 157 - 161 (47, 90)

A.H. GREENAWAY, J.C. DAINTY [1978] "On long-baseline amplitude interferometers in astronomical applications" Optica Acta, Vol. 25 No. 3, March. Pp. 181 - 189 (85, 86)

A. GRENVIK, U. HEDSTRAND, H. SJÖGREN [1966] "Problems in Pneumotachography" Acta Anaesthesiologica Scandinavia, Vol. 10 No. 3. Pp. 147 - 155 (276)

H.R. GRIBBIN [1983] "Using body surface movements to study breathing" Journal of Medical Engineering & Technology, Vol. 7 No. 5, September/October. Pp. 217 - 223 (284, 285, 296)

C. GUILLEMINAULT, A. TILKIAN, W.C. DEMENT [1976] "The sleep apnea syndromes" in "Annual Review of Medicine" Vol. 27, W.P. Creger, C.H. Coggins, E.W. Hancock (eds), Annual Reviews Inc., Palo Alto, California. Pp. 465 - 484 (346)

S.F. GULL, G.J. DANIELL [1978] "Image reconstruction from incomplete and noisy data" Nature, Vol. 272 No. 5655, 20 April. Pp. 686 - 690 (121)

A.C. GUYTON [1971] "Basic Human Physiology: Normal Function and Mechanisms of Disease" W.B. Saunders, Philadelphia. (270, 271, 272, 322, 348)

R.W. HAMMING [1973] "Numerical Methods for Scientists and Engineers", 2nd edition McGraw-Hill, New York. (180)

R. HANBURY BROWN [1974] "The Intensity Interferometer: Its Application to Astronomy" Taylor & Francis, London. (83, 84, 85, 108, 209, 214)

R. HANBURY BROWN [1984] "Measuring the sizes of stars"
Journal of Astrophysics and Astronomy, Vol. 5 No. 1, March.
Pp. 19 - 30 (85)

R. HANBURY BROWN, R.Q. TWISS [1954] "A new type of interferometer for use in radio astronomy"
Philosophical Magazine, Vol. 45 No. 366, July. Pp. 663 - 682 (85)

R. HANBURY BROWN, J. DAVIS, L.R. ALLEN [1974] "The angular diameters of 32 stars"
Monthly Notices of the Royal Astronomical Society, Vol. 167 No. 1, April. Pp. 121 - 136 (85)

J.W. HARDY [1978] "Active optics: a new technology for the control of light"
Proceedings of the IEEE, Vol. 66 No. 6, June. Pp. 651 - 697 (83)

F.J. HARRIS [1978] "On the use of windows for harmonic analysis with the discrete Fourier transform"
Proceedings of the IEEE, Vol. 66 No. 1, January. Pp. 51 - 83 (381)

M.H. HAYES [1984] "Signal reconstruction from spectral phase or spectral magnitude"
Chapter 3 in "Image Reconstruction from Incomplete Observations", T.S. Huang (ed.), JAI Press, Greenwich, Connecticut. Advances in Computer Graphics and Image Processing, Vol. 1. Pp. 145 - 189 (130)

S. HAYKIN [1978] "Communication Systems"
John Wiley & Sons, New York. (67)

J.N. HEASLEY [1984] "Numerical restoration of astronomical images"
Publications of the Astronomical Society of the Pacific, Vol. 96 No. 583, September. Pp. 767 - 772 (83)

J.C. HEBDEN, E.K. HEGE, J.M. BECKERS [1985] "Differential speckle imaging with the cophased Multiple Mirror Telescope"
in "International Conference on Speckle", H.H. Arsenault (ed.), Proceedings of the SPIE, Vol. 556, 20 - 23 August. Pp. 284 - 289 (90)

E.K. HEGE [1984] Private communication to R.H.T. Bates
Steward Observatory, University of Arizona, Tucson, 15 September. (209, 210)

E.K. HEGE, E.N. HUBBARD, J.D. DRUMMOND, P.A. STRITTMATTER, S.P. WORDEN, T. LAUER [1982a] "Speckle interferometric observations of Pluto and Charon"
Icarus, Vol. 50 No. 1, April. Pp. 72 - 81 (100)

E.K. HEGE, E.N. HUBBARD, P.A. STRITTMATTER, W.J. COCKE [1982b] "The Steward Observatory speckle interferometry system"
Optica Acta, Vol. 29 No. 5, May. Pp. 701 - 715 (108, 209, 227)

E.K. HEGE, J.M. BECKERS, P.A. STRITTMATTER, D.W. McCARTHY [1985] "Multiple Mirror Telescope as a phased array telescope"
Applied Optics, Vol. 24 No. 16, 15 August. Pp. 2565 - 2576 (90, 210, 217, 379)

G. HENDENSTIERNA, A. STRANDBERG, B. BRISMAR, H. LUNDQUIST, L. SVENSSON, L. TOKICS [1985] "Functional residual capacity, thoracoabdominal dimensions, and central blood volume during general anesthesia with muscle paralysis and mechanical ventilation" Anesthesiology, Vol. 62 No. 3, March. Pp. 247 - 254 (278)

D.W. HILL [1973] "Electronic Techniques in Anaesthesia and Surgery", 2nd edition
Butterworths, London. (275, 276, 280, 299, 303, 346)

S.L. HILL, J.P. BLACKBURN, T.R. WILLIAMS [1982] "Measurement of respiratory flow by inductance pneumography"
Medical & Biological Engineering & Computing, Vol. 20 No. 4, July. Pp. 517 - 518 (282, 315)

K.-H. HOFMANN, G. WEIGELT [198_] "Imaging speckle interferometer in space (ISIS): image reconstruction by speckle masking"
Journal of the Optical Society of America A,
submitted for publication October 1985. (140)

K.-H. HOFMANN, R. LADEBECK, G. WEIGELT [198_] "Speckle masking of the central object in NGC 3603"
Astronomy and Astrophysics, submitted for publication 1985. (140)

J.A. HÖGBOM [1974] "Aperture synthesis with a non-regular distribution of interferometer baselines"
Astronomy and Astrophysics, Supplement Series, Vol. 15 No. 3, June. Pp. 417 - 426 (156)

R. R. HOWELL, M.T. McGINN [1985] "Infrared speckle observations of Io: an eruption in the Loki region"
Science, Vol. 230 No. 4721, 4 October. Pp. 63 - 65 (129)

R.H. HUDGIN [1977] "Wave-front reconstruction for compensated imaging"
Journal of the Optical Society of America, Vol. 67 No. 3, March. Pp. 375 - 378 (129)

R.E. HUFNAGEL [1974] "Variations of Atmospheric Turbulence"
Digest of Technical Papers, Topical Meeting on Optical Propagation through Turbulence, University of Colorado, Boulder, Colorado. 9 - 11 July. Pp. WA1/1 - WA1/4 (not sighted) (54)

R.E. HUFNAGEL, N.R. STANLEY [1964] "Modulation transfer function associated with image transmission through turbulent media"
Journal of the Optical Society of America, Vol. 54 No. 1, January. Pp. 52 - 61 (60)

A.M.J. HUISER [1982a] "Astro speckle interferometry: a simple modification resolving the phase problem"
Optik, Vol. 61 No. 4, July. Pp. 433 - 436 (131)

A.M.J. HUISER [1982b] "A procedure to correct the images of astronomical objects for the distortions due to atmospheric turbulence"
Optics Communications, Vol. 42 No. 4, 15 July. Pp. 226 - 230 (90)

R. HULL, W.G. VAN AKEN, J. HIRSH, A.S. GALLUS, G. HOICKA, A.G.G. TURPIE, I. WALKER, M. GENT [1976] "Impedance plethysmography using the occlusive cuff technique in the diagnosis of venous thrombosis"
Circulation, Vol. 53 No. 4, April. Pp. 696 - 700 (373)

R. HULL, D.W. TAYLOR, J. HIRSH, D.L. SACKETT, P. POWERS, A.G.G. TURPIE, I. WALKER [1978] "Impedance plethysmography: the relationship between venous filling and sensitivity and specificity for proximal vein thrombosis"
Circulation, Vol. 58 No. 5, November. Pp. 898 - 902 (373)

B.R. HUNT [1979] "Matrix formulation of the reconstruction of phase values from phase differences"
Journal of the Optical Society of America, Vol. 69 No. 3, March. Pp. 393 - 399 (129)

B.R. HUNT, W.R. FRIGHT, R.H.T. BATES [1983] "Analysis of the shift-and-add method for imaging through turbulent media"
Journal of the Optical Society of America, Vol. 73 No. 4, April. Pp. 456 - 465 (68, 153, 156, 157, 162, 165, 378, 380)

D.J. HUTTER, H.A. McALISTER, W.I. HARTKOPF [1985] "Accurate differential magnitudes of binary star components as obtained using the SAA algorithm" (conference abstract)
Bulletin of the American Astronomical Society, Vol. 17 No. 2. P. 551 (158, 377)

INTEL [1978] "PL/M-86 Programming Manual", O/N 9800466A
Intel Corporation, Santa Clara, California. (357)

INTEL [1981] "iSBC 88/25 Single Board Computer Hardware Reference Manual", O/N 143825-001
Intel Corporation, Santa Clara, California. (356)

J.B. IRWIN [1975] "Report from Gainesville"
Sky and Telescope, Vol. 49 No. 3, March. Pp. 164 - 166 (64, 150)

A. ISHIMARU [1978] "Wave Propagation and Scattering in Random Media", Volumes 1 and 2
Academic Press, New York. (28, 32, 33, 57)

K. ITOH, Y. OHTSUKA [1983] "Phase estimation based on the maximum likelihood criterion"
Applied Optics, Vol. 22 No. 19, 1 October. Pp. 3054 - 3057 (124)

A. ITOH, A. ISHIDA, N. KIKUCHI, N. OKAZAKI, T. ISHIHARA, S. KIRA [1982a] "Non-invasive ventilatory volume monitor"
Medical & Biological Engineering & Computing, Vol. 20 No. 5, September. Pp. 613 - 619 (280)

A. ITOH, A. ISHIDA, N. KIKUCHI, N. OKAZAKI, Y. KURATOMI, T. ISHIHARA, S. KIRA [1982b] "Microcomputer-based respiratory function monitoring system using impedance pneumography"
Medical & Biological Engineering & Computing, Vol. 20 No. 5, September. Pp. 620 - 624 (281, 290)

P. JACQUINOT, B. ROIZEN-DOSSIER [1964] "Apodisation"
in "Progress in Optics" Vol. III, E. Wolf (ed.), North-Holland, Amsterdam. Pp. 29 - 186 (70)

J.H. JAFFE, W.R. MARTIN [1980] "Opioid analgesics and antagonists"
Chapter 22 in "The Pharmacological Basis of Therapeutics", 6th edition, A. Goodman Gilman, L.S. Goodman, A. Gilman (eds.), Macmillan, New York. Pp. 494 - 534 (294, 295)

M.Y. JAFFRIN, C. VANHOUTTE [1979] "Quantitative interpretation of arterial impedance plethysmographic signals"
Medical & Biological Engineering & Computing, Vol. 17 No. 1, January. Pp. 2 - 10 (350)

M.A. JENKINS [1975] "Algorithm 493. Zeros of a real polynomial [C2]"
ACM Transactions on Mathematical Software, Vol. 1 No. 2, June. Pp. 178 - 189 (235)

M.A. JENKINS, J.F. TRAUB [1972] "Algorithm 419 - zeros of a complex polynomial [C2]"
Communications of the ACM, Vol. 15 No. 2, February. Pp. 97 - 99 (235)

R.C. JENNISON [1958] "A phase sensitive interferometer technique for the measurement of the Fourier transforms of spatial brightness distributions of small angular extent"
Monthly Notices of the Royal Astronomical Society, Vol. 118 No. 3. Pp. 276 - 284 (125)

D.S. JONES [1964] "The Theory of Electromagnetism"
Pergamon Press, Oxford. (20, 21)

C. JORDAN [1982] "Assessment of the effects of drugs on respiration"
British Journal of Anaesthesia, Vol. 54 No. 7, July. Pp. 763 - 782 (295, 296, 320)

E.C. JORDAN, K.G. BALMAIN [1968] "Electromagnetic Waves and Radiating Systems", 2nd edition
Prentice-Hall, Englewood Cliffs, New Jersey (18)

C. JORDAN, J.G. JONES, D.J. PINTO [1980] "Postoperative respiration monitoring"
in "Proceedings of the Third International Symposium on Ambulatory Monitoring (ISAM 1979)", F.D. Stott, E.B. Raftery, L. Goulding (eds), Academic Press, London. Pp. 373 - ? (not sighted) (282)

D.P. KARO, A.M. SCHNEIDERMAN [1975] "Image reconstruction in speckle interferometry"
in "Imaging in Astronomy", Technical Digest from the Topical Meeting on Imaging in Astronomy, Cambridge, Massachusetts, 18 - 21 June. Pp. ThC4-1 - ThC4-3 (129, 217, 378)

D.P. KARO, A.M. SCHNEIDERMAN [1976a] "Nonstationarity of the atmospheric transfer function"
Journal of the Optical Society of America, Vol. 66 No. 10, October. Pp. 1065 - 1066 (62)

D.P. KARO, A.M. SCHNEIDERMAN [1976b] "Speckle interferometry lens-atmosphere MTF measurements"
atmospheric transfer function"
Journal of the Optical Society of America, Vol. 66 No. 11, November.
Pp. 1252 - 1256 (71)

D.P. KARO, A.M. SCHNEIDERMAN [1977] "Speckle interferometry with severely aberrated telescopes"
Journal of the Optical Society of America, Vol. 67 No. 9, September.
Pp. 1277 - 1278 (71)

D.P. KARO, A.M. SCHNEIDERMAN [1978] "Speckle interferometry at finite spectral bandwidths and exposure times"
Journal of the Optical Society of America, Vol. 68 No. 4, April.
Pp. 480 - 485 (61, 66)

I.R. KING [1971] "The profile of a star image"
Publications of the Astronomical Society of the Pacific, Vol. 83
No. 492, April. Pp. 199 - 201 (63)

S. KIRA, Y. HUKUSHIMA, S. KITAMURA, A. ITO [1971] "Transthoracic electrical impedance variations associated with respiration"
Journal of Applied Physiology, Vol. 30 No. 6, June. Pp. 820 - 826
(280)

K.T. KNOX [1976] "Image retrieval from astronomical speckle patterns"
Journal of the Optical Society of America, Vol. 66 No. 11, November.
Pp. 1236 - 1239 (125, 126, 127, 129)

K.T. KNOX, B.J. THOMPSON [1973] "New methods of processing speckle pattern star images"
The Astrophysical Journal, Vol. 182 No. 3 Part 2, 15 June.
Pp. L133 - L136 (125)

K.T. KNOX, B.J. THOMPSON [1974] "Recovery of images from atmospherically degraded short-exposure photographs"
The Astrophysical Journal, Vol. 193 No. 1 Part 2, 1 October.
Pp. L45 - L48 (125, 126, 128)

D.E. KNUTH [1969] "The Art of Computer Programming", Volumes 1 - 3
Addison-Wesley, Massachusetts. (181)

K. KONNO, J. MEAD [1967] "Measurement of the separate volume changes of rib cage and abdomen during breathing"
Journal of Applied Physiology, Vol. 22 No. 3, March. Pp. 407 - 422
(279, 283, 285, 288)

D. KORFF [1973] "Analysis of a method for obtaining near-diffraction-limited information in the presence of atmospheric turbulence"
Journal of the Optical Society of America, Vol. 63 No. 8, August.
Pp. 971 - 980 (57, 71, 76, 101)

D. KORFF, G. DRYDEN, M.G. MILLER [1972] "Information retrieval from atmospheric induced speckle patterns"
Optics Communications, Vol. 5 No. 3, June. Pp. 187 - 192 (58, 70)

D. KORFF, G. DRYDEN, R.P. LEAVITT [1975] "Isoplanicity: the translation invariance of the atmospheric Green's function" Journal of the Optical Society of America, Vol. 65 No. 11, November. Pp. 1321 - 1330 (77, 81, 198)

E. KREYSZIG [1979] "Advanced Engineering Mathematics", 4th edition John Wiley & Sons, New York. (20, 22, 31, 124)

A. LABEYRIE [1970] "Attainment of diffraction limited resolution in large telescopes by Fourier analysing speckle patterns in star images" Astronomy and Astrophysics, Vol. 6 No. 1, May (II). Pp. 85 - 87 (90, 94, 96, 123)

A. LABEYRIE [1975] "Interference fringes obtained on Vega with two optical telescopes" The Astrophysical Journal, Vol. 196 No. 2 Part 2, 1 March. Pp. L71 - L75 (83)

A. LABEYRIE [1976] "High-resolution techniques in optical astronomy" in "Progress in Optics" Vol. XIV, E. Wolf (ed.), North-Holland, Amsterdam. Pp. 47 - 87 (62, 79, 91, 176)

A. LABEYRIE [1978] "Stellar interferometry methods" in "Annual Review of Astronomy and Astrophysics", Vol. 16, Annual Reviews Inc., Palo Alto, California. Pp. 77 - 102 (91, 149)

A. LABEYRIE [1982] "Stellar interferometry: a widening frontier" Sky & Telescope, Vol. 63 No. 4, April. Pp. 334 - 338 (84)

R.G. LANE, W.R. FRIGHT, R.H.T. BATES [198_] "Direct phase retrieval" IEEE Transactions on Acoustics, Speech, and Signal Processing, submitted for publication December 1985. (159)

S. LAXMINARAYAN, O. MILLS, L. MICHELSON, A.C. CORNWELL, A. MARMAROU, E.F. COSTIGAN Jr, E.D. WEITZMAN [1983] "Sudden infant death syndrome: a digital computer-based apnoea monitor" Medical & Biological Engineering & Computing, Vol. 21 No. 2, March. Pp. 191 - 196 (273)

J.C. LEADER [1978] "The generalized partial coherence of a radiation source and its far-field" Optica Acta, Vol. 25 No. 5, May. Pp. 395 - 413 (43)

B.Ja. LEVIN [1964] "Distribution of Zeros of Entire Functions" (translated from the Russian) American Mathematical Society, Providence, Rhode Island. (232)

R.M. LEWITT, R.H.T. BATES [1978] "Image reconstruction from projections: I: general theoretical considerations" Optik, Vol. 50 No. 1, February. Pp. 19 - 33 (140, 266)

R.M. LEWITT, R.H.T. BATES, T.M. PETERS [1978] "Image reconstruction from projections: II: modified back-projection methods" Optik, Vol. 50 No. 2, March. Pp. 85 - 109 (140, 266)

E.H. LINFOOT, R.C. WITCOMB [1972] "Random wave-front perturbations and telescopic star images"
Monthly Notices of the Royal Astronomical Society, Vol. 158 No. 2.
Pp. 199 - 231 (82)

L.T. LITTLE, A. HEWISH [1966] "Interplanetary scintillation and its relation to the angular structure of radio sources"
Monthly Notices of the Royal Astronomical Society, Vol. 134 No. 3.
Pp. 221 - 237 (58)

C.Y.C. LIU, A.W. LOHMANN [1973] "High resolution image formation through the turbulent atmosphere"
Optics Communications, Vol. 8 No. 4, August. Pp. 372 - 377 (73, 113)

A.W. LOHMANN, G.P. WEIGELT [1975] "Large field speckle interferometry"
in "Imaging in Astronomy", Technical Digest from the Topical Meeting on Imaging in Astronomy, Cambridge, Massachusetts, 18 - 21 June.
Pp. ThC3-1 - ThC3-4 (114)

A.W. LOHMANN, G.P. WEIGELT [1977] "Image reconstruction from astronomical speckle interferograms"
in "Optical Telescopes of the Future", F. Pacini, W. Richter, R.N. Wilson (eds.),
Proceedings of ESO Conference, Geneva, 12 - 15 December.
Pp. 479 - 494 (114, 121)

A.W. LOHMANN, G.P. WEIGELT [1979] "Astronomical speckle interferometry; measurements of isoplanicity and of temporal correlation"
Optik, Vol. 53 No. 3, June. Pp. 167 - 180 (66, 74, 78, 79, 81, 202)

A.W. LOHMANN, B. WIRNITZER [1984] "Triple correlations"
Proceedings of the IEEE, Vol. 72 No. 7, July. Pp. 889 - 901 (131, 140)

A.W. LOHMANN, G. WEIGELT, B. WIRNITZER [1983] "Speckle masking in astronomy: triple correlation theory and applications"
Applied Optics, Vol. 22 No. 24, 15 December. Pp. 4028 - 4037 (131, 133, 135, 136, 137, 138, 140)

J.S. LUNDSGAARD, J. GRØNLUND, N. EINER-JENSEN [1979] "Evaluation of a constant-temperature hot-wire anenometer for respiratory-gas-flow measurements"
Medical & Biological Engineering & Computing, Vol. 17 No. 2, March.
Pp. 211 - 215 (277)

C.R. LYNDS, S.P. WORDEN, J.W. HARVEY [1974]
Preprints of the 144th American Astronomical Society Meeting, Gainesville, Florida, 10 - 13 December. (not sighted) (150)

C.R. LYNDS, S.P. WORDEN, J.W. HARVEY [1976] "Digital image reconstruction applied to Alpha Orionis"
The Astrophysical Journal, Vol. 207 No. 1 Part 1, 1 July.
Pp. 174 - 180, Plates 3, 4 (146, 147, 149, 150, 151, 164)

J.M. MARIOTTI, A. CHELLI, R. FOY, P. LENA, F. SIBILLE, G. TCHOUNTONOV [1983] "Infrared speckle imaging: improvement of the method; results on Miras and protostars"
Astronomy and Astrophysics, Vol. 120 No. 2, April (II). Pp. 237 - 248 (66, 103)

J.C. MAXWELL [1865] "A Dynamical Theory of the Electromagnetic Field"
Philosophical Magazine, Fourth Series, Vol. 29, January - June.
Pp. 152 - 157 (19)

H.A. McALISTER [1977] "Binary-star speckle interferometry"
Sky and Telescope, Vol. 53 No. 5, May. Pp. 346 - 350 (91, 100)

H.A. McALISTER, W.I. HARTKOPF [1984] "Catalog of Interferometric Measurements of Binary Stars"
CHARA Contribution No. 1, January. Center for High Angular Resolution Astronomy, Georgia State University, Atlanta, Georgia. (100, 183)

D.W. McCARTHY Jr, R.G. PROBST, F.J. LOW [1985] "Infrared detection of a close cool companion to van Biesbroeck 8"
The Astrophysical Journal, Vol. 290 No. 1 Part 2, 1 March.
Pp. L9 - L13 (100)

M.J. McDONNELL, R.H.T. BATES [1975] "Preprocessing of degraded images to augment existing restoration methods"
Computer Graphics and Image Processing, Vol. 4 No. 1, March 1975.
Pp. 25 - 39 (381)

M.J. McDONNELL, R.H.T. BATES [1976] "Digital restoration of an image of Betelgeuse"
The Astrophysical Journal, Vol. 208 No. 2 Part 1, 1 September.
Pp. 443 - 452 (149, 150, 151)

B.L. McGLAMERY [1967] "Restoration of turbulence-degraded images"
Journal of the Optical Society of America, Vol. 57 No. 3, March.
Pp. 293 - 297 (83)

B.L. McGLAMERY [1971] "Image restoration techniques applied to astronomical photography"
in "Astronomical Use of Television Type Image Sensors", NASA Technical Report No. SP-256. Pp. 167 - 192 (not sighted) (123)

A.E. McKINNON, M.J. McDONNELL, P.J. NAPIER, R.H.T. BATES [1976] "Self-consistent deconvolution: II: applications"
Optik, Vol. 44 No. 3, February. Pp. 253 - 272 (229)

J. MEAD, N. PETERSON, G. GRIMBY, J. MEAD [1967] "Pulmonary ventilation measured from body surface movements"
Science, Vol. 156 No. 3780, 9 June. Pp. 1383 - 1384 (279, 284)

J.H. MEIJER, I.P.H. REUBEN, H. SCHNEIDEN, P.L. OE, W. ALLON, L.G. THIJS [1982] "Differential impedance plethysmography for measuring thoracic impedances"
Medical & Biological Engineering & Computing, Vol. 20 No. 2, March.
Pp. 187 - 194 (280)

E.B. MERRICK, T.J. HAYES [1976] "Continuous, non-invasive measurements of arterial blood oxygen levels"
Hewlett-Packard Journal, Vol. 28 No. 2, October. Pp. 2 - 9 (296)

L.N. MERTZ [1979] "Speckle imaging, photon by photon"
Applied Optics, Vol. 18 No. 5, 1 March. Pp. 611 - 614 (123, 124)

L.N. MERTZ [1984] "Phase estimation with few photons"
Applied Optics, Vol. 23 No. 10, 15 May. Pp. 1638 - 1641 (124, 146)

A.A. MICHELSON, F.G. PEASE [1921] "Measurement of the diameter of α Orionis with the interferometer"
The Astrophysical Journal, Vol. 53 No. 4, May. Pp. 249 - 259 (83, 84)

J.S. MILLEDGE, F.D. STOTT [1977] "Inductive plethysmography - a new respiratory transducer"
The Journal of Physiology, Vol. 267 No. 1, May. Pp. 4P - 5P (280, 281, 343)

M.G. MILLER, A.M. SCHNEIDERMAN, P.F. KELLEN [1973] "Comparison of methods for processing short-exposure data from large telescopes"
The Astrophysical Journal, Vol. 186 No. 2 Part 2, 1 December. Pp. L91 - L93 (125)

M.O. MILNER [1979] "Error Correction in Images and Imaging Instruments"
Ph.D. thesis, University of Canterbury, Christchurch. (151)

M.O. MILNER, R.H.T. BATES [1979] "Imaging despite seeing"
New Zealand Journal of Science, Vol. 22 No. 4, December. Pp. 579 - 580 (151)

R.A. MINARD [1985] "Imaging in a Distorting Medium"
Ph.D. thesis, University of Canterbury, Christchurch. (69, 160, 161)

R.A. MINARD, B.S. ROBINSON, R.H.T. BATES [1985] "Full-wave computed tomography part 3: coherent shift-and-add imaging"
IEE Proceedings A, Vol. 132 No. 1, January. Pp. 50 - 58 (68, 160, 161, 162)

S.N. MOHAPATRA, H.M. ARENSON [1979] "The measurement of peripheral blood flow by the electrical impedance technique"
Journal of Medical Engineering & Technology, Vol. 3 No. 3, May. Pp. 132 - 137 (350)

P.M. MORSE, H. FESHBACH [1953] "Methods of Theoretical Physics" Parts I, II
McGraw-Hill, New York. (30, 31)

P.M. MORSE, K.U. INGARD [1968] "Theoretical Acoustics"
McGraw-Hill, New York. (16)

P.J. NAPIER [1971] "Reconstruction of Radiating Sources"
Ph.D. thesis, University of Canterbury, Christchurch. (229, 230, 232)

P.J. NAPIER, R.H.T. BATES [1971] "Holographic approach to radiation pattern measurement - II. Experimental verification"
International Journal of Engineering Science, Vol. 9 No. 12, December. Pp. 1193 - 1208 (111)

R. NARAYAN, R. NITYANANDA [198_] "Maximum entropy image restoration in astronomy"
To appear in "Annual Review of Astronomy and Astrophysics" Vol. 24, 1986. (121)

J. NEWSOM DAVIS, D. STAGG [1975] "Interrelationships of the volume and time components of individual breaths in resting man"
The Journal of Physiology, Vol. 245 No. 2, February. Pp. 481 - 498 (272)

I. NEWTON [1704] "Opticks: or, a Treatise of the Reflexions, Refractions, Inflexions and Colours of Light"
Printers to the Royal Society, Prince's Arms, St. Paul's Churchyard, London. (37)

M. NIETO-VESPERINAS, J.C. DAINTY [1984] "Testing for uniqueness of phase recovery in two dimensions"
Optics Communications, Vol. 52 No. 2, 15 November. Pp. 94 - 98 (118)

P. NISENSON, C. PAPALIOLIOS [1983] "Effects of photon noise on speckle image reconstruction with the Knox-Thompson algorithm"
Optics Communications, Vol. 47 No. 2, 15 August. Pp. 91 - 96 (108, 109, 124, 129, 227)

P. NISENSON, R.V. STACHNIK [1975] "Real-time speckle imaging"
in "Imaging in Astronomy", Technical Digest from Topical Meeting on Imaging in Astronomy, Cambridge, Massachusetts, 18 - 21 June. Pp. ThC2-1 - ThC2-4 (129)

P. NISENSON, R.V. STACHNIK [1978] "Measurements of atmospheric isoplanatism using speckle interferometry"
Journal of the Optical Society of America, Vol. 68 No. 2, February. Pp. 169 - 175 (79)

P. NISENSON, R. STACHNIK, C. PAPALIOLIOS [1980] "High resolution imaging at large telescopes"
in "Optical & Infrared Telescopes for the 1990s", A. Hewitt (ed.), Proceedings of KPNO Conference, Tucson, Arizona, 7 - 12 January. Pp. 401 - 417 (129)

P. NISENSON, J. APT, R. GOODY, P. HOROWITZ [1981] "Radius and limb darkening of Titan from speckle imaging"
The Astronomical Journal, Vol. 86 No. 11, November. Pp. 1690 - 1693 (129)

P. NISENSON, J. APT, R. GOODY, C. PAPALIOLIOS [1983a] "Speckle imaging for planetary research"
Icarus, Vol. 53 No. 3, March. Pp. 465 - 478 (91, 102, 129)

P. NISENSON, R.V. STACHNIK, C. PAPALIOLIOS, P. HOROWITZ [1983b] "Data recording and processing for speckle image reconstruction" in "Applications of Speckle Phenomena", W.H. Carter (ed.), Proceedings of the SPIE, Vol. 243, 29 - 30 July. Pp. 88 - 94 (128, 129)

P. NISENSON, R.V. STACHNIK, M. KAROVSKA, R. NOYES [1985] "A new optical source associated with T Tauri" The Astrophysical Journal, Vol. 297 No. 1 Part 2, 1 October. Pp. L17 - L20, Plate L1 (129)

R. NITYANANDA, R. NARAYAN [1982] "Maximum entropy image reconstruction - a practical non-information-theoretic approach" Journal of Astrophysics and Astronomy, Vol. 3 No. 4, December. Pp. 419 - 450 (121)

D.E. NOVOSELLER [1985] "Upper bound on turbulence-induced log amplitude fluctuations" Applied Optics, Vol. 24 No. 1, 1 January. P. 10 (75)

R.W. NOYES, R.V. STACHNIK, P. NISENSON [1981] "Speckle image reconstruction of solar features" Report No. AFGL-TR-81-0155 to Air Force Geophysics Laboratory, U.S. Air Force, Hanscom AFB, Massachusetts, 15 May. (129)

J.F. NUNN [1977] "Applied Respiratory Physiology", 2nd edition Butterworths, London. (270, 271, 272, 273, 275, 276, 277, 295)

J. NYBOER [1970] "Electrical Impedance Plethysmography", 2nd edition C.C. Thomas, Springfield, Illinois. (280, 350)

B.T. O'CONNOR, T.S. HUANG [1981] "Application of phase unwrapping to image restoration" Computer Graphics and Image Processing, Vol. 15 No. 1, January. Pp. 25 - 44 (123, 124, 129)

K.A. O'DONNELL, B.J. BRAMES, J.C. DAINTY [1982] "Measurements of the spatial-temporal statistics of stellar speckle patterns at Mauna Kea, Hawaii" Optics Communications, Vol. 41 No. 2, 15 March. Pp. 79 - 82 (66)

J. OHTSUBO [1985] "Effects of finite spectral bandwidth and focusing error on the transfer function in stellar speckle interferometry" Journal of the Optical Society of America A, Vol. 2 No. 5, May. Pp. 667 - 673 (82)

A.V. OPPENHEIM, J.S. LIM [1981] "The importance of phase in signals" Proceedings of the IEEE, Vol. 69 No. 5, May. Pp. 529 - 541 (130)

A.V. OPPENHEIM, M.H. HAYES, J.S. LIM [1982] "Iterative procedures for signal reconstruction from Fourier transform phase" Optical Engineering, Vol. 21 No. 1, January/February. Pp. 122 - 127 (130)

J.J. OSBORN [1977] "Cardiopulmonary monitoring in the respiratory intensive care unit" Medical Instrumentation, Vol. 11 No. 5, September-October. Pp. 278 - 282 (273, 277)

J.J. OSBORN [1978] "A flowmeter for respiratory monitoring"
Critical Care Medicine, Vol. 6 No. 5, September/October.
Pp. 349 - 351 (277)

C.M. OWEN, M.M. SMITH, G. BOWES [1985] "Effect of changing rib cage / abdominal contributions on the accuracy of respiratory inductive plethysmography" (conference abstract)
Volume, Vol. 5 No. 2, June. P. 9 (282, 289)

C. PAPALIOLIOS, P. NISENSON, S. EBSTEIN [1985] "Speckle imaging with the PAPA detector"
Applied Optics, Vol 24 No. 2, 15 January. Pp. 287 - 292 (90, 129)

G. PARRY, J.G. WALKER, R.J. SCADDAN [1979] "On the statistics of stellar speckle patterns and pupil plane scintillation"
Optica Acta, Vol. 26 No. 5, May. Pp. 563 - 574 (66)

R. PETROV, F. RODDIER, C. AIME [198_] "The signal-to-noise ratio in differential speckle interferometry"
Journal of the Optical Society of America A,
submitted for publication June 1984. (90)

S. POLLAINE, A. BUFFINGTON, F.S. CRAWFORD [1979] "Measurement of the size of the isoplanatic patch using a phase-correcting telescope"
Journal of the Optical Society of America, Vol. 69 No. 1, January.
Pp. 84 - 89 (79, 83)

K.J.D. QUINN [1974] "Eton Four-figure Mathematical and Statistical Tables", 4th edition
Eton Press, Christchurch. (328)

N. RAWAL, M. WATTWIL [1984] "Respiratory depression after epidural morphine - an experimental and clinical study"
Anesthesia and Analgesia, Vol. 63 No. 1, January. Pp. 8 - 14 (295)

A.A.G. REQUICHA [1980] "The zeros of entire functions: theory and engineering applications"
Proceedings of the IEEE, Vol. 68 No. 3, March. Pp. 308 - 328 (230, 234, 235)

W.T. RHODES, J.W. GOODMAN [1973] "Interferometric technique for recording and restoring images degraded by unknown aberrations"
Journal of the Optical Society of America, Vol. 63 No. 6, June.
Pp. 647 - 657 (47)

E. RIBAK [198_] "Astronomical imaging by filtered, weighted shift-and-add"
Journal of the Optical Society of America A,
submitted for publication January 1986. (162, 164)

E. RIBAK, E. LEIBOWITZ [1985] "Shearing stellar interferometer. 1: digital data analysis scheme"
Applied Optics, Vol. 24 No. 18, 15 September. Pp. 3088 - 3093 (85)

E. RIBAK, E.K. HEGE, J.C. CHRISTOU [1984] "Identification of speckles by matched filtering" (conference abstract)
Bulletin of the American Astronomical Society, Vol. 16 No. 4.
Pp. 885 - 886 (162, 164)

E. RIBAK, E.K. HEGE, J.C. CHRISTOU [1985a] "Use of matched filtering to identify speckle locations"
in "International Conference on Speckle", H.H. Arsenault (ed.), Proceedings of the SPIE, Vol. 556, 20 - 23 August. Pp. 196 - 201 (162, 164)

E. RIBAK, E. LEIBOWITZ, E.K. HEGE [1985b] "Shearing stellar interferometer. 2: optoelectronic phase-locked system"
Applied Optics, Vol. 24 No. 18, 15 September. Pp. 3094 - 3100 (85)

G. RICORT, C. AIME [1979] "Solar seeing and the statistical properties of the photospheric solar granulation. III. Solar speckle interferometry"
Astronomy and Astrophysics, Vol. 76 No. 3, July (II). Pp. 324 - 335 (52, 66, 71, 100, 101)

C.H. ROBERTSON Jr, M.E. BRADLEY, L.D. HOMER [1980] "Comparison of two- and four-magnetometer methods of measuring ventilation"
Journal of Applied Physiology, Vol. 49 No. 3, September.
Pp. 355 - 362 (279)

B.S. ROBINSON [1982] "Speckle Processing for Ultrasonic Imaging"
Ph.D. thesis, University of Canterbury, Christchurch. (174)

B.S. ROBINSON, R.H.T. BATES [1980] "Wideband ultrasonic diffraction measurements"
Australasian Physical & Engineering Sciences in Medicine, Vol. 3 No. 6, November/December. Pp. 233 - 238 (159)

C. RODDIER [1976] "Measurements of the atmospheric attenuation of the spectral components of astronomical images"
Journal of the Optical Society of America, Vol. 66 No. 5, May.
Pp. 478 - 482 (61, 62, 85, 86)

F. RODDIER [1981] "The effects of atmospheric turbulence in optical astronomy"
in "Progress in Optics" Vol. XIX, E. Wolf (ed.), North-Holland, Amsterdam. Pp. 281 - 376 (47, 49, 52, 54, 55, 56, 57, 58, 59, 60, 61, 63, 66, 82, 83)

F. RODDIER [1982] "How to achieve diffraction limited resolution with large space telescopes"
in "Advanced Space Instrumentation in Astronomy", R.M. Bonnet (ed.), Pergamon Press, Oxford, 1983.
Advances in Space Research, Vol. 2 No. 4. Pp. 3 - 9 (86)

F. RODDIER [1986] Private communication to R.H.T. Bates
National Optical Astronomy Observatories, Tucson, Arizona, 20 January. (121)

F. RODDIER, J.B. BRECKINRIDGE [1984] "Interferometric image reconstruction using the L.D.R. in a light bucket mode"
Bulletin of the American Astronomical Society, Vol. 16 No. 3 Part II.
Pp. 832 - 837 (86)

C. RODDIER, F. RODDIER [1973] "Correlation measurements on the complex amplitude of stellar plane waves perturbed by atmospheric turbulence"
Journal of the Optical Society of America, Vol. 63 No. 6, June. Pp. 661 - 663 (61)

C. RODDIER, F. RODDIER [1978] "Imaging with a coherence interferometer in optical astronomy"
in "Image Formation from Coherence Functions in Astronomy", C. van Schooneveld (ed.), D. Reidel Publishing Company, Dordrecht, Holland, 1979. Proceedings of IAU Colloquium No. 49, Groningen, Netherlands, 10 - 12 August. Pp. 175 - 185 (83, 86)

C. RODDIER, F. RODDIER [1983] "High angular resolution observations of Alpha Orionis with a rotation shearing interferometer"
The Astrophysical Journal, Vol. 270 No. 1 Part 2, 1 July. Pp. L23 - L26 (86)

F. RODDIER, C. RODDIER [1985] "An image reconstruction of Alpha Orionis"
The Astrophysical Journal, Vol. 295 No. 1 Part 2, 1 August. Pp. L21 - L23, Plate L2 (86, 121, 158, 163)

F. RODDIER, G. RICORT, C. RODDIER [1978] "Defocusing effects in astronomical speckle interferometry"
Optics Communications, Vol. 24 No. 3, March. Pp. 281 - 284 (71)

F. RODDIER, J.M. GILLI, J. VERNIN [1982] "On the isoplanatic patch size in stellar speckle interferometry"
Journal of Optics (Paris), Vol. 13 No. 2, March-April. Pp. 63 - 70 (58, 66, 74, 75, 76, 77, 78, 79, 196, 197, 200, 202, 216, 217, 378)

G.L. ROGERS [1979] "The stellar interferometry of a star cluster with a prominent variable"
Optics Communications, Vol. 30 No. 1, July. Pp. 1 - 4 (114)

N. ROTH [1978] ""First stammering of the heart": Ludwig's kymograph"
Medical Instrumentation, Vol. 12 No. 6, November-December. P. 348 (274)

D. ROYSTON, C. JORDAN, J.G. JONES [1981] "Effect of nitrous oxide on the adaptation to inspiratory resistive loading in man"
British Journal of Anaesthesia, Vol. 53 No. 6, June. Pp. 667P - 668P (282)

G. RUPPEL [1982] "Manual of Pulmonary Function Testing", 3rd edition C.V. Mosby, St. Louis, Missouri. (272, 274, 275, 277, 278, 296)

F.D. RUSSELL, J.W. GOODMAN [1971] "Nonredundant arrays and postdetection processing for aberration compensation in incoherent imaging"
Journal of the Optical Society of America, Vol. 61 No. 2, February. Pp. 182 - 191 (47)

J.D. SACKNER, M.J. BROUDY, B. DAVIS, M.A. COHN, M.A. SACKNER [1980a] "Ventilation at rest and during exercise measured without physical connection to the airway"
in "Proceedings of the Third International Symposium on Ambulatory Monitoring (ISAM 1979)", F.D. Stott, E.B. Raftery, L. Goulding (eds), Academic Press, London. Pp. 341 - 353 (282)

J.D. SACKNER, A.J. NIXON, B. DAVIS, N. ATKINS, M.A. SACKNER [1980b] "Non-invasive measurement of ventilation during exercise using a respiratory inductive plethysmograph. I."
American Review of Respiratory Disease, Vol. 122 No. 6, December. Pp. 867 - 871 (280, 281, 282, 315, 343)

J.D. SACKNER, A.J. NIXON, B. DAVIS, N. ATKINS, M.A. SACKNER [1980c] "Effects of breathing through external dead space on ventilation at rest and during exercise. II."
American Review of Respiratory Disease, Vol. 122 No. 6, December. Pp. 933 - 940 (278, 282)

M.A. SACKNER, H.F. GONZALEZ, G. JENOURI, M. RODRIGUEZ [1984a] "Effects of abdominal and thoracic breathing on breathing pattern components in normal subjects and in patients with chronic obstructive pulmonary disease"
American Review of Respiratory Disease, Vol. 130 No. 4, October. Pp. 584 - 587 (282)

M.A. SACKNER, H. GONZALEZ, M. RODRIGUEZ, A. BELSITO, D.R. SACKNER, S. GRENVIK [1984b] "Assessment of asynchronous and paradoxic motion between rib cage and abdomen in normal subjects and in patients with chronic obstructive pulmonary disease"
American Review of Respiratory Disease, Vol. 130 No. 4, October. Pp. 588 - 593 (282, 288, 300, 326, 330, 346)

A.V. SAHAKIAN, W.J. TOMPKINS, J.G. WEBSTER [1985] "Electrode motion artifacts in electrical impedance pneumography"
IEEE Transactions on Biomedical Engineering, Vol. BME-32 No. 6, June. Pp. 448 - 451 (280)

R.J. SCADDAN, J.C. DAINTY [1977] "A simple method of estimating the RMS phase variation due to atmospheric turbulence"
Optics Communications, Vol. 21 No. 1, April. Pp. 51 - 54 (56)

A.M. SCHNEIDERMAN, D.P. KARO [1978a] "Speckle interferometry measurements of atmospheric nonisoplanicity using double stars"
Journal of the Optical Society of America, Vol. 68 No. 3, March. Pp. 338 - 347 (79, 81, 198)

A.M. SCHNEIDERMAN, D.P. KARO [1978b] "Measurement of r_o with speckle interferometry"
Journal of the Optical Society of America, Vol. 68 No. 3, March. Pp. 348 - 351 (62)

A.D. SEAGAR [1983a] "Probing With Low Frequency Electric Currents"
Ph.D. thesis, University of Canterbury, Christchurch. (-15, -13, 280, 299, 347, 349, 350, 362, 366)

A.D. SEAGAR [1983b] Private communications
University of Canterbury, Christchurch. (364, 365, 366)

A.D. SEAGAR, J.M. GIBBS, F.M. DAVIS [1984] "Interpretation of venous occlusion plethysmographic measurements using a simple model"
Medical & Biological Engineering & Computing, Vol. 22 No. 1, January.
Pp. 12 - 18 (-13, 299, 350, 351, 353, 354, 355, 366, 384)

S. SEVITT, N. GALLAGHER [1961] "Venous thrombosis and pulmonary embolism. A clinico-pathological study in injured and burned patients"
British Journal of Surgery, Vol. XLVIII No. 211, March. Pp. 475 - 489 (349)

M. SHULMAN, A.N. SANDLER, J.W. BRADLEY, P.S. YOUNG, J. BREBNER [1984] "Postthoracotomy pain and pulmonary function following epidural and systemic morphine"
Anesthesiology, Vol. 61 No. 5, November. Pp. 569 - 575 (296)

F. SIBILLE, A. CHELLI, P. LENA [1979] "Infrared speckle interferometry"
Astronomy and Astrophysics, Vol. 79 No. 3, November (I).
Pp. 315 - 328 (100, 105, 169)

S. SILVER (ed.) [1965] "Microwave Antenna Theory and Design"
Dover Publications, New York. (11, 12, 27)

J.J. SINTON [1979] "An 8080 microprocessor software development system"
Final year project report, Department of Electrical and Electronic Engineering, University of Canterbury. (356)

A.M. SINTON, B.L.K. DAVEY, R.H.T. BATES [198_] "Augmenting shift-and-add with zero-and-add"
Journal of the Optical Society of America A, to appear June 1986.
(144, 145, 160, 161, 162, 229, 265, 266)

H.H. SKILLING [1974] "Electric Networks"
John Wiley & Sons, New York. (310)

R.R. SOKAL, F.J. ROHLF [1981] "Biometry. The Principles and Practice of Statistics in Biological Research"
W.H. Freeman, San Francisco. (315)

J.L. SPENCER, B.A. ZIKRIA, J.M. KINNEY, J.R. BROELL, T.M. MICHAILOFF, A.B. LEE [1972] "A system for continuous measurement of gas exchange and respiratory functions"
Journal of Applied Physiology, Vol. 33 No. 4, October. Pp. 523 - 528 (275, 326)

R.V. STACHNIK, P. NISENSON, D.C. EHN, R.H. HUDGIN, V.E. SCHIRF [1977] "Speckle image reconstruction of solar features"
Nature, Vol. 266 No. 5598, 10 March. Pp. 149 - 151 (129)

R.V. STACHNIK, P. NISENSON, R.W. NOYES [1983] "Speckle image reconstruction of solar features"
The Astrophysical Journal, Vol. 271 No. 1 Part 2, 1 August.
Pp. L37 - L40, Plates L2 - L4 (129)

D. STAGG, M. GOLDMAN, J. NEWSOM DAVIS [1978] "Computer-aided measurement of breath volume and time components using magnetometers" Journal of Applied Physiology, Vol. 44 No 4, April. Pp. 623 - 633 (285, 286)

W.D. STANLEY [1975] "Digital Signal Processing" Reston Publishing Company, Reston, Virginia. (234)

A.R. STARK [1982] "Apnea monitors - 1982" Medical Instrumentation, Vol. 16 No. 3, May-June. Pp. 160 - 162 (277, 279)

R.G. STEWART [1983] "The good news - P696, P796 now IEEE standards" IEEE Micro, Vol. 3 No. 1, February. Pp. 69 - 71 (318)

P.A. STRITTMATTER [1977] "Multiple mirror telescopes" in "Optical Telescopes of the Future", F. Pacini, W. Richter, R.N. Wilson (eds.), Proceedings of ESO Conference, Geneva, 12 - 15 December. Pp. 165 - 184 (214)

J.W. STROHBEHN [1971] "Optical propagation through the turbulent atmosphere" in "Progress in Optics" Vol. IX, E. Wolf (ed.), North-Holland, Amsterdam. Pp. 75 - 122 (52)

J. STRONG [1958] "Concepts of Classical Optics" W.H. Freeman & Co., San Francisco. (19)

D.S. SUMNER [1978a] "Volume plethysmography in vascular disease: an overview" in "Noninvasive Diagnostic Techniques in Vascular Disease", E.F. Bernstein (ed.), C.V. Mosby, St. Louis, Illinois. Pp. 68 - 92 (349, 353)

D.S. SUMNER [1978b] "Mercury strain-gauge plethysmography" in "Noninvasive Diagnostic Techniques in Vascular Disease", E.F. Bernstein (ed.), C.V. Mosby, St. Louis, Illinois. Pp. 126 - 147 (280, 350)

R. SUNTHERALINGAM [1985a] "Non-invasive breath monitoring" Final year project progress report, Department of Electrical and Electronic Engineering, University of Canterbury. (309)

R. SUNTHERALINGAM [1985b] "Non-invasive breath monitoring" Final year project final report, Department of Electrical and Electronic Engineering, University of Canterbury. (383)

M.K. SYKES, M.D. VICKERS, C.J. HULL [1981] "Principles of Clinical Measurement", 2nd edition Blackwell Scientific Publications, Oxford. (275, 276, 278)

SYSTEMS GROUP [1981] "The System 2800 family of business computers" (sales brochure) Systems Group, Orange, California. (318)

E. TABACHNIK, N. MULLER, B. TOYE, H. LEVISON [1981] "Measurement of ventilation in children using the respiratory inductive plethysmograph"
The Journal of Pediatrics, Vol. 99 No. 6, December. Pp. 895 - 899 (278, 282, 289)

W.J. TANGO [1979] "A note on two aperture pupil plane stellar speckle interferometry"
Optica Acta, Vol. 26 No. 1, January. Pp. 109 - 111 (85)

W.J. TANGO, R.Q. TWISS [1980] "Michelson stellar interferometry"
in "Progress in Optics" Vol. XVII, E. Wolf (ed.), North-Holland, Amsterdam. Pp. 239 - 277 (80)

V.I. TATARSKI [1961] "Wave Propagation in a Turbulent Medium"
(translated from the Russian)
McGraw-Hill, New York. (57, 58)

K. THEORELL, H.F.R. PRECHTL, J.E. VOS [1974] "A polygraphic study of normal and abnormal newborn infants"
Neuropädiatrie, Vol. 5 No. 3, August. Pp. 279 - 317 (331)

D.H. THORPE [1984] "Opiate structure and activity - a guide to understanding the receptor"
Anesthesia and Analgesia, Vol. 63 No. 2, February. Pp. 143 - 151 (294)

G.M. TISI [1980] "Pulmonary Physiology in Clinical Medicine"
Williams & Wilkins, Baltimore. (272, 274, 275, 312)

A. TITLE, T. POPE, S. SCHOOLMAN [1975] "The temporal and spatial extent of the isoplanatic patch" (conference abstract)
Bulletin of the American Astronomical Society, Vol. 7 No. 3 Part I. P. 462 (79)

M.J. TOBIN, M.A. SACKNER [1982] "Monitoring smoking patterns of low and high tar cigarettes with inductive plethysmography"
American Review of Respiratory Disease, Vol. 126 No. 2, August. Pp. 258 - 264 (282)

M.J. TOBIN, G. JENOURI, B. LIND, H. WATSON, A. SCHNEIDER, M.A. SACKNER [1983a] "Validation of respiratory inductive plethysmography in patients with pulmonary disease"
Chest, Vol. 83 No. 4, April. Pp. 615 - 620 (282, 289, 331)

M.J. TOBIN, T.S. CHADHA, G. JENOURI, S.J. BIRCH, H.B. GAZEROGLU, M.A. SACKNER [1983b] "Breathing patterns 1. Normal subjects"
Chest, Vol. 84 No. 2, August. Pp. 202 - 205 (282)

M.J. TOBIN, T.S. CHADHA, G. JENOURI, S.J. BIRCH, H.B. GAZEROGLU, M.A. SACKNER [1983c] "Breathing patterns 2. Diseased subjects"
Chest, Vol. 84 No. 3, September. Pp. 286 - 294 (282)

J.M. TRIBOLET [1977] "A new phase unwrapping algorithm"
IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-25 No. 2, April. Pp. 170 - 177 (123)

D.G. TUCKER [1967] "Sonar in Fisheries - A Forward Look"
Fishing News (Books), London. (336)

S.Z. TURNEY, W. BLUMENFELD [1973] "Heated Fleisch pneumotachometer: a calibration procedure"
Journal of Applied Physiology, Vol. 34 No. 1, January. Pp. 117 - 121 (277)

P. VAN TOORN, A.H. GREENAWAY, A.M.J. HUISER [1984] "Phaseless object reconstruction"
Optica Acta, Vol. 31 No. 7, July. Pp. 767 - 774 (118)

K. VON DER HEIDE [1978] "Image restoration by the method of least squares"
Astronomy and Astrophysics, Vol. 70 No. 6, December (II).
Pp. 777 - 784 (151)

O. VON DER LÜHE [1984] "Estimating Fried's parameter from a time series of an arbitrary resolved object imaged through atmospheric turbulence"
Journal of the Optical Society of America A, Vol. 1 No. 5, May.
Pp. 510 - 519 (62, 101)

J.G. WALKER [1981a] "The phase retrieval problem. A solution based on zero location by exponential apodisation"
Optica Acta, Vol. 28 No. 6, June. Pp. 735 - 738 (141, 142, 143)

J.G. WALKER [1981b] "Object reconstruction from turbulence-degraded images"
Optica Acta, Vol. 28 No. 8, August. Pp. 1017 - 1019 (141, 143, 229)

J.G. WALKER [1982] "Computer simulation of a method for object reconstruction from stellar speckle interferometry data"
Applied Optics, Vol. 21 No. 17, 1 September. Pp. 3132 - 3137 (144)

P.G.M. WALLACE, W. NORRIS [1975] "The management of postoperative pain"
British Journal of Anaesthesia, Vol. 47 No. 2, February.
Pp. 113 - 120 (294)

C.P. WANG [1975] "Isoplanicity for imaging through turbulent media"
Optics Communications, Vol. 14 No. 2, June. Pp. 200 - 204 (73)

R.H. WARREN, S.H. ALDERSON [1984] "The assessment of wheezing in infants and young children utilizing respiratory inductive plethysmography: a case report"
Journal of the Arkansas Medical Society, Vol. 81 No. 2, July.
Pp. 133 - 136 (not sighted) (282)

R.H. WARREN, S.H. ALDERSON [1985] "Calibration of computer-assisted (Respicomp) respiratory inductive plethysmography in newborns"
American Review of Respiratory Disease, Vol. 131 No. 4, April.
Pp. 564 - 567 (282, 285, 289)

H. WATSON [1979] "The technology of respiratory inductive plethysmography"
in "Proceedings of the Third International Symposium on Ambulatory Monitoring (ISAM 1979)", F.D. Stott, E.B. Raftery, L. Goulding (eds), Academic Press, London. Pp. 537 - 563 (281, 315, 343)

H.L. WATSON, S.J. BIRCH, W.G. GRUEN, M.A. COHN, A.W. SCHNEIDER, M.A. SACKNER [1981] "Comparison of calibration techniques for respiratory inductive plethysmograph in humans and sheep"
American Review of Respiratory Disease (annual meeting supplement), Vol. 123 No. 4 Part 2, April. P. 181 (282, 289)

G.P. WEIGELT [1975] "Large field speckle interferometry"
Optik, Vol. 43 No. 2, August. Pp. 111 - 128 (79, 114)

G.P. WEIGELT [1977] "Modified astronomical speckle interferometry 'Speckle masking'"
Optics Communications, Vol. 21 No. 1, April. Pp. 55 - 59 (131, 135, 136)

G.P. WEIGELT [1978] " Speckle holography measurements of the stars Zeta Cancri and ADS 3358"
Applied Optics, Vol. 17 No. 17, 1 September. Pp. 2660 - 2662 (112, 112)

G.P. WEIGELT [1979] "High-resolution astrophotography. New isoplanicity measurements and speckle holography applications"
Optica Acta, Vol. 26 No. 11, November. Pp. 1351 - 1357 (79, 112)

G.P. WEIGELT [1980] "Stellar speckle interferometry and speckle holography at low light levels"
in "Applications of Speckle Phenomena", W.H. Carter (ed.), Proceedings of the SPIE, Vol. 243, 29 - 30 July. Pp. 103 - 111 (112)

G.P. WEIGELT, G. BAIER [1985] "R136a in the 30 Doradus nebula resolved by holographic speckle interferometry"
Astronomy and Astrophysics, Vol. 150 No. 1, September (I). Pp. L18 - L20 (112)

G. WEIGELT, B. WIRNITZER [1983] "Image reconstruction by the speckle-masking method"
Optics Letters, Vol. 8 No. 7, July. Pp. 389 - 391 (131, 135, 140)

G.P. WEIGELT, G. BAIER, R. LADEBECK [1985a] "R136a and the central object in the giant HII region NGC 3603 resolved by holographic speckle interferometry"
The Messenger, No. 40, June. Pp. 4 - 6 (112)

G.P. WEIGELT, G. BAIER, J. EBERSBERGER, F. FLEISCHMANN, K.-H. HOFMANN, R. LADEBECK [1985b] "Speckle interferometry, image reconstruction by speckle masking, speckle spectroscopy, multiple-mirror interferometry"
in "International Conference on Speckle", H.H. Arsenault (ed.), Proceedings of the SPIE, Vol. 556, 20 - 23 August. Pp. 238 - 247 (90, 140)

J.V. WEIL, R.E. McCULLOUGH, J.S. KLINE, I.E. SODAL [1975] "Diminished ventilatory response to hypoxia and hypercapnia after morphine in normal man"
The New England Journal of Medicine, Vol. 292 No. 21, 22 May.
Pp. 1103 - 1106 (295)

C. WEISSMAN, J. ASKANAZI, J. MILIC-EMILI, J.M. KINNEY [1984] "Effect of respiratory apparatus on respiration"
Journal of Applied Physiology, Vol. 57 No. 2, August. Pp. 475 - 480 (278)

W.T. WELFORD [1976] "Aplanatism and isoplanatism"
in "Progress in Optics" Vol. XIII, E. Wolf (ed.), North-Holland, Amsterdam. Pp. 267 - 292 (5)

G.L. WELTER, S.P. WORDEN [1978] "A method for processing stellar speckle interferometry data"
Journal of the Optical Society of America, Vol. 68 No. 9, September.
Pp. 1271 - 1275 (101, 137, 156)

G.L. WELTER, S.P. WORDEN [1980] "The angular diameters of supergiant stars from speckle interferometry"
The Astrophysical Journal, Vol. 242 No. 2 Part 1, 1 December.
Pp. 673 - 683 (149, 150)

H.B. WHEELER [1978] "Impedance plethysmography: theoretical and experimental basis"
in "Noninvasive Diagnostic Techniques in Vascular Disease",
E.F. Bernstein (ed.), C.V. Mosby, St. Louis, Illinois. Pp. 112 - 125 (350)

H.B. WHEELER, F.A. ANDERSON Jr, J.M. MATESANZ, J.E. LARSEN [1978] "Impedance phlebography: the diagnosis of venous thrombosis by occlusive impedance plethysmography"
in "Noninvasive Diagnostic Techniques in Vascular Disease",
E.F. Bernstein (ed.), C.V. Mosby, St. Louis, Illinois. Pp. 359 - 373 (373)

R.L. WILDEY [1973] "Spatial filtering of astronomical photographs: III. Remarks on the Knox-Thompson reapodization"
The Astrophysical Journal, Vol. 186 No. 1 Part 2, 15 November.
Pp. L47 - L50 (125)

M.S. WILKERSON, S.P. WORDEN [1977] "Further speckle interferometric studies of α Orionis"
The Astronomical Journal, Vol. 82 No. 8, August. Pp. 642 - 645 (150)

A.J. WILSON, B.H. BROWN [1982] "Wideband compander system for biological signals"
Medical & Biological Engineering & Computing, Vol. 20 No. 4, July.
Pp. 512 - 516 (344)

A.J. WILSON, C.I. FRANKS [1982] "The Sheffield respiration analysis system"
IEE Proceedings A, Vol. 129 No. 9, December. Pp. 702 - 706 (281, 304, 326, 327, 331, 346)

A.J. WILSON, C.I. FRANKS, I.L. FREESTON [1982a] "Algorithms for the detection of breaths from respiratory waveform recordings of infants"
Medical & Biological Engineering & Computing, Vol. 20 No. 3, May.
Pp. 286 - 292 (326, 327)

A.J. WILSON, C.I. FRANKS, I.L. FREESTON [1982b] "Methods of filtering the heart-beat artefact from the breathing waveform of infants obtained by impedance pneumography"
Medical & Biological Engineering & Computing, Vol. 20 No. 3, May.
Pp. 293 - 298 (280, 300)

B. WIRNITZER [1985] "Bispectral analysis at low light levels and astronomical speckle masking"
Journal of the Optical Society of America A, Vol. 2 No. 1, January.
Pp. 14 - 21 (130, 134, 141)

N. WIRTH [1976] "Algorithms + Data Structures = Programs"
Prentice-Hall, Englewood Cliffs, New Jersey. (170)

D.H. WITHERS [1974] "Remark on algorithm 419 [C2]"
Communications of the ACM, Vol. 17 No. 3, March. P. 157 (235)

M.C. WON, D. MNYAMA, R.H.T. BATES [1985] "Improving initial phase estimates for phase retrieval algorithms"
Optica Acta, Vol. 32 No. 4, April. Pp. 377 - 396 (119)

N.J. WOOLF [1982] "High resolution imaging from the ground"
in "Annual Review of Astronomy and Astrophysics" Vol. 20.
G. Burbidge (ed.), Annual Reviews Inc., Palo Alto, California.
Pp. 367 - 398 (52, 54, 62, 63, 64)

S.P. WORDEN [1975] "The angular diameter of Alpha Herculis A"
The Astrophysical Journal, Vol. 201 No. 2 Part 2, 15 October.
Pp. L69 - L70 (150)

S.P. WORDEN [1976] "Digital analysis of speckle photographs: the angular diameter of Arcturus"
Publications of the Astronomical Society of the Pacific, Vol. 88
No. 521, February. Pp. 69 - 72 (150)

S.P. WORDEN [1977] "Astronomical image reconstruction"
Vistas in Astronomy, Vol. 20 Part 3. Pp. 301 - 318 (83, 91)

S.P. WORDEN, M.K. STEIN [1979] "Angular diameters of the asteroids Vesta and Pallas determined from speckle observations"
The Astronomical Journal, Vol. 84 No. 1, January. Pp. 140 - 142
(102)

S.P. WORDEN, C.R. LYNDS, J.W. HARVEY [1976] "Reconstructed images of Alpha Orionis using stellar speckle interferometry"
Journal of the Optical Society of America, Vol. 66 No. 11, November.
Pp. 1243 - 1246 (147, 149, 150)

S.P. WORDEN, M.K. STEIN, G.D. SCHMIDT, J.R.P. ANGEL [1977] "The angular diameter of Vesta from speckle interferometry"
Icarus, Vol. 32 No. 4, December. Pp. 450 - 457 (101)

B.M. WRIGHT [1955] "A respiratory anenometer"
The Journal of Physiology, Vol. 127 No. 2, 28 February. P. 25P (275, 277)

B.M. WRIGHT [1977] "An abdominal respiration detector"
The Journal of Physiology, Vol. 271 No. 2, October. Pp. 11P - 12P (279)

C.G. WYNNE [1979] "Extending the bandwidth of speckle interferometry"
Optics Communications, Vol. 28 No. 1, January. Pp. 21 - 25 (67)

A.T. YOUNG [1974] "Seeing: its cause and cure"
The Astrophysical Journal, Vol. 189 No. 3 Part 1, 1 May.
Pp. 587 - 604 (56, 82)

R. ZAKS [1980] "The CP/M Handbook with MP/M"
Sybex, USA. (319)

P.V. ZIMMERMAN, S.J. CONNELLAN, H.C. MIDDLETON, M.V. TABONA, M.D. GOLDMAN, N. PRIDE [1983] "Postural changes in rib cage and abdominal volume-motion coefficients and their effect on the calibration of a respiratory inductance plethysmograph"
American Review of Respiratory Disease, Vol. 127 No. 2, February.
Pp. 209 - 214 (282, 284, 288, 289, 303)